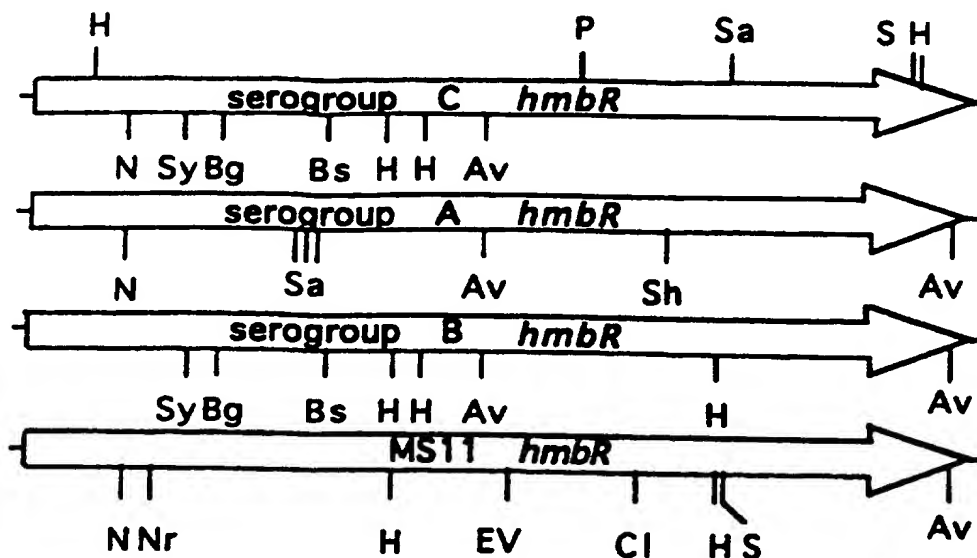




## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12N 15/12, C07K 14/22, 16/12, A61K 38/16, 39/095, G01N 33/68</b>		A2	(11) International Publication Number: <b>WO 96/12020</b>
			(43) International Publication Date: 25 April 1996 (25.04.96)
(21) International Application Number: PCT/US95/13623		(US). NASSIF, Xavier [FR/FR]; 36. rue Miollis, F-75015 Paris (FR).	
(22) International Filing Date: 17 October 1995 (17.10.95)		(74) Agent: NOONAN, Kevin, E.; Banner & Allegretti, Ltd., Ten South Wacker Drive, Chicago, IL 60606 (US).	
(30) Priority Data: 08/326,670 18 October 1994 (18.10.94) US Not furnished 2 October 1995 (02.10.95) US		(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).	
(60) Parent Application or Grant (63) Related by Continuation US Filed on 08/326,670 (CIP) 18 October 1994 (18.10.94)		Published Without international search report and to be republished upon receipt of that report.	
(71) Applicant (for all designated States except US): OREGON HEALTH SCIENCES UNIVERSITY [US/US]; 3181 S.W. Sam Jackson Park Road, Portland, OR 97201-3098 (US).			
(72) Inventors; and (75) Inventors/Applicants (for US only): STOJILJKOVIC, Igor [HR/US]; 3223 S.W. 11th Avenue #3, Portland, OR 97201 (US). SO, Magdalene [US/US]; 777 S.W. 48th Drive, Portland, OR 97221 (US). HWA, Vivian [AU/US]; 7011 S.W. 4th Avenue, Portland, OR 97219 (US). HEFFRON, Fred [US/US]; 17887 Hillside Drive, West Linn, OR 97068			

(54) Title: HEMOGLOBIN RECEPTORS FROM NEISSERIAE



## (57) Abstract

The present invention relates to novel bacterial hemoglobin receptor proteins and genes that encode such proteins. The invention is directed toward the isolation, characterization, diagnostic and therapeutic use of bacterial hemoglobin receptor proteins, nucleic acids encoding such proteins, recombinant expression constructs comprising such nucleic acids and cells transformed therewith, and antibodies and epitopes of such hemoglobin receptor proteins. The invention relates particularly to hemoglobin receptor proteins and genes encoding such proteins from *Neisseria* species, especially *N. meningitidis* and serotypes thereof, and *N. gonorrhoeae*. Methods for the diagnostic and therapeutic use of the proteins, epitopes, antibodies and nucleic acids of the invention are also provided, including the use of the proteins, epitopes, antibodies and nucleic acids of the invention for the production of vaccines effective in providing immunization of a human against infection by pathogenic bacteria of *Neisseria* species.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LJ	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

## HEMOGLOBIN RECEPTORS FROM NEISSERIAE

This invention was made with government support under National Institute of Health grants R01 AI32493 and R01 AI22933. The U.S. government has certain rights to this invention.

## BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to hemoglobin receptor genes and the proteins encoded therefrom of certain bacterial species, particularly species of *Neisseria* bacteria. More particularly, this invention relates to hemoglobin receptor genes, polypeptides and peptides useful for preparing vaccines and antibodies against *Neisseria*, and methods and means for producing such peptides and polypeptides *in vitro*. Also provided are diagnostic and therapeutic methods and reagents useful in detecting and treating *Neisseria* infection and methods for developing novel and effective anti-*Neisseria* agents.

2. Background of the Invention

The *Neisseriae* comprise a genus of bacteria that includes two gram-negative species of pyogenic cocci pathogenic for humans: *Neisseria meningitidis* and *Neisseria gonorrhoeae*. *N. meningitidis* is a major cause of bacterial meningitis in humans, especially children. The disease characteristically proceeds from asymptomatic carriage of the bacterium in the nasopharynx to invasion of the bloodstream and cerebrospinal fluid in susceptible individuals.

*Neisseria meningitidis* is one of the leading causes of bacterial meningitis in children and healthy adults in the world. The severity of the disease is evidenced by the ability of meningococci to cause the death of previously healthy individuals in less than 24 hours. *N. meningitidis* has a polysaccharide capsule whose diversity of component antigenic polysaccharide molecules has resulted in the classification of ten different serogroups. Of these, group A strains are the classic epidemic strains; group B and C are generally endemic strains, but C occasionally causes an epidemic outbreak. All known group A strains have the same protein antigens on their outer membranes, while group B strains have a dozen serotypes or groupings based on the

presence of principal outer membrane protein antigens (as opposed to polysaccharides).

Survival of a pathogen such as *N. meningitidis* in a host depends on its ability to overcome a battery of host defense mechanisms. One nonspecific host defense mechanism against microbial intruders is to limit the availability of iron in tissues (Weinberg, 1984, *Physiological. Rev.* 64: 65-102), because iron is a necessary nutrient for most microbial pathogens. The vast majority of iron in the human adult is located intracellularly in the form of hemoglobin (76%) or ferritin (23%). The remainder can be found extracellularly bound to host iron-binding proteins such as transferrin and lactoferrin (Otto *et al.*, 1992, *Crit. Rev. Microbiol.* 18: 217-233).

Pathogenic bacteria have adapted to this iron-limiting environment by developing highly specific and effective iron assimilation systems. A large number of these bacteria secrete siderophores, small, non-protein iron chelators which, due to their extremely high affinity for iron (III), scavenge trace amounts of iron(III) from the environment and shuttle the iron back to the bacterial cell (Baggs and Neilands, 1987, *Microbiol. Rev.* 51: 509-518; Braun and Hantke, 1991, in Winkelmann (ed.), *Handbook of Microbial Iron Chelates*, CRC Press: Boca Raton, Fla., pp. 107-138.).

Alternatively, some bacterial pathogens, like *Neisseriae* species (Archibald and DeVoe, 1979, *FEMS Microbiol. Lett.* 6: 159-162; Mickelson *et al.*, 1982, *Infect. Immun.* 35: 915-920; Dyer *et al.*, 1987, *Infect. Immun.* 55: 2171-2175), *Haemophilus influenzae* (Coulton and Pang, 1983, *Curr. Microbiol.* 9: 93-98; Schryvers, 1988, *Mol. Microbiol.* 2: 467-472; Jarosik *et al.*, 1994, *Infect. Immun.* 62: 2470-2477), *Vibrio cholerae* (Stoebner and Payne, 1988, *Infect. Immun.* 56: 2891-2895; Henderson and Payne, 1994, *J. Bacteriol.* 176: 3269-3277), *Yersinia* (Stojiljkovic and Hantke, 1992, *EMBO J.* 11: 4359-4367) and *Actinobacillus pleuropneumoniae* (Gerlach *et al.*, 1992, *Infect. Immun.* 60: 3253-3261) have evolved more sophisticated mechanisms to sequester iron from the host. These pathogens can directly bind host's iron-binding proteins such as lactoferrin, transferrin, and heme-containing compounds, and use them as sole sources of iron.

The importance of iron in the virulence of *N. meningitidis* was demonstrated by *in vivo* studies using mice as the animal model system (Calver *et al.*, 1976, *Can.*



*J. Microbiol.* 22: 832-838; Holbien *et al.*, 1981, *Infect. Immun.* 34: 120-125). Specific iron-regulated outer membrane receptors have been shown to be involved in the binding and the utilization of lactoferrin- and transferrin-iron in *Neisseriae* (Schryvers and Morris, 1988, *Infect. Immun.* 56: 1144-1149 and *Mol. Microbiol.* 2: 281-288; Legrain *et al.*, 1993, *Gene* 130: 81-90; Pettersson *et al.*, 1993, *Infect. Immun.* 61: 4724-4733 and 1994, *J. Bacteriol.* 176: 1764-1766). These receptors share significant amino acid similarity and, most probably, also the mechanism of iron internalization, with receptors for siderophores and vitamin B12 of other Gram-negative bacteria (Cornelissen *et al.*, 1993, *J. Bacteriol.* 174: 5788-5797). In contrast, the mechanism by which *Neisseriae* utilize hemoglobin- and hemin-iron as well as the components involved have so far not been described.

Recently, several proteins with hemoglobin-binding and/or hemin-binding activities have been identified in total membranes of iron-limited *N. meningitidis* and *N. gonorrhoeae*.

Lee and Hill, 1992, *J. gen. Microbiol.* 138: 2647-2656 disclose the specific hemoglobin binding by isolated outer membranes of *N. meningitidis*.

Martek and Lee, 1994, *Infect. Immun.* 62: 700-703 disclosed that acquisition of heme iron by *N. meningitidis* does not involve meningococcal transferrin-binding proteins.

Lee, 1994, *Microbiol.* 140: 1473-1480 describes the biochemical isolation and characterization of hemin binding proteins from *N. meningitidis*.

The precise role of these proteins in hemin and/or hemoglobin utilization remains unclear at present, although these proteins are likely to be components of a hemin-utilization system in *N. meningitidis*.

The dependence on host iron stores for *Neisseria* growth is a potentially useful route towards the development of novel and effective therapeutic intervention strategies. Historically, infections of both *N. meningitidis* and *N. gonorrhoeae* were treated chemoprophylactically with sulfonamide drugs. However, with the development of sulfonamide-resistant strains came the necessity of using alternative modes of therapy such as antibiotic treatment. More recently, the drug treatment of choice includes the administration of high grade penicillin. However, the success of antimicrobial treatment is decreased if therapy is not initiated early after infection.

Gonococcal infection has also been treated with penicillin, ampicillin, or amoxicillin, tetracycline hydrochloride, and spectinomycin. Unfortunately, because the incidence of infections due to penicillinase-producing bacteria has increased, several new, more expensive  $\beta$ -lactam antibiotics have been used in treatment. Despite the fact that existing antibiotics have decreased the serious consequences of gonorrhea, their use has not lowered the incidence of the infection in the general population.

Prevention of meningococcal disease has been attempted by chemoprophylaxis and immunoprophylaxis. At present, rifampin and minocycline are used, but only for humans in close contact with an infected person as this treatment has a number of disadvantages. The only commercially available vaccine against meningococcal meningitis has as its major component the bacterial polysaccharide capsule. In adults this vaccine protects against serogroups A, C, Y and W135. It is not effective against serogroup B, and is ineffective in children against serogroup C. Thus far, immunoprophylactic preventive treatment has not been available for *N. gonorrhoeae*.

Thus, what is needed are better preventative therapies for meningococcal meningitis and gonorrhea including more effective, longer lasting vaccines which protect across all of the serogroups of *N. meningitidis* and all the serotypes of *N. gonorrhoeae*. In addition, better methods are need to treat meningococcal and gonococcal infection.

### SUMMARY OF THE INVENTION

The present invention relates to the cloning, expression and functional characterization of genes encoding bacterial hemoglobin receptor proteins. Specifically, the invention relates to genes encoding hemoglobin receptor proteins from *Neisseria* species, in particular *Neisseria meningitidis* and *N. gonorrhoeae*. The invention comprises species of nucleic acids having a nucleotide sequence encoding novel bacterial hemoglobin receptor proteins. Also provided by this invention is the deduced amino acid sequence of the cognate hemoglobin receptor proteins of these bacterial genes.

The invention provides nucleic acids, nucleic acid hybridization probes, recombinant expression constructs capable of expressing the hemoglobin receptor

protein of the invention in cultures of transformed cells, preferably bacterial cells, and such cultures of transformed bacterial cells that express the hemoglobin receptor proteins of the invention. The invention also provides gene knockout vectors for inactivating the hemoglobin receptor protein gene in cells, particularly cells of *Neisseria* species, *via*, for example, homologous recombination and other mechanisms, and cultures of such hemoglobin receptor protein null mutant cells.

The invention also provides homogeneous preparations of the bacterial hemoglobin receptor proteins of the invention, as well as antibodies against and epitopes of the hemoglobin receptor protein. Methods for characterizing this receptor protein and methods for using the protein in the development of agents having pharmacological uses related to this receptor, particularly bactericidal and bacteriostatic uses, are also provided by the invention.

In other embodiments of this invention are provided diagnostic methods and reagents encompassing the use of the anti-*Neisseria* hemoglobin receptor protein antibodies of the invention. Still further embodiments provided herein include therapeutic methods and reagents encompassing the use of the anti-*Neisseria* hemoglobin receptor protein antibodies of the invention. Even more embodiments include diagnostic methods and reagents encompassing the use of the *Neisseria* hemoglobin receptor protein-encoding nucleic acids of the invention, as sensitive probes for the presence of *Neisseria* infection using nucleic acid hybridization techniques and/or *in vitro* amplification methodologies. Yet additional embodiments of the invention include therapeutic methods and reagents encompassing the use of the *Neisseria* hemoglobin receptor protein-encoding nucleic acids of the invention, comprising recombinant expression constructs engineered to produce antisense transcripts of the *Neisseria* hemoglobin receptor gene and fragments thereof, as well as recombinant knockout vectors of the invention. The invention also provides the *Neisseria* hemoglobin receptor protein and epitopes thereof as components of vaccines for the development of non-disease associated immunity to pathological infection with bacteria of *Neisseria* species.

In a first aspect, the invention provides a nucleic acid having a nucleotide sequence encoding a bacterial hemoglobin receptor protein gene. In a preferred embodiment, the bacterial hemoglobin receptor protein gene is isolated from bacteria

of *Neisseria* species. In a particularly preferred embodiment, the hemoglobin receptor protein gene is isolated from *Neisseria meningitidis*, serotype C. In a particular example of this embodiment, the nucleic acid comprises a 3.3 kilobase (kb) *Bam*HI/*Hind*III fragment of *N. meningitidis* genomic DNA. In this embodiment, the nucleotide sequence comprises an open reading frame of 2376 nucleotides of *N. meningitidis* genomic DNA encoding 792 amino acids comprising the hemoglobin receptor gene. In this embodiment of the invention, the nucleotide sequence of the *N. meningitidis* hemoglobin receptor gene is the sequence depicted in Figure 2 (SEQ ID No.:1). It will be understood that the *N. meningitidis* gene as disclosed herein is defined, insofar as is necessary, by the amino acid sequence of the protein encoded therein, said amino acid sequence being represented in Figure 2 (SEQ. ID No.:2). Thus, it will be understood that the particular nucleotide sequence depicted in Figure 2 (SEQ. ID. No.:1) is but one of a number of equivalent nucleotide sequences that encode the hemoglobin receptor protein, due to the degeneracy of the genetic code, and that all such alternative, equivalent nucleotide sequences are hereby explicitly encompassed within the disclosed nucleotide sequences of the invention. Also included herein are any mutant or allelic variations of this nucleotide sequence, either naturally occurring or the product of *in vitro* chemical or genetic modification. Each such variant will be understood to have essentially the same nucleotide sequence as the nucleotide sequence of the corresponding *N. meningitidis* hemoglobin receptor protein disclosed herein.

In another particularly preferred embodiment of this aspect of the invention, the hemoglobin receptor protein gene is isolated from *Neisseria meningitidis*, serotype A. In a particular example of this embodiment, the nucleic acid comprises a 2373 basepair (bp) polymerase chain reaction-amplified fragment of *N. meningitidis*, serotype A genomic DNA. In this embodiment, the nucleotide sequence comprises an open reading frame of 2373 nucleotides of *N. meningitidis* genomic DNA encoding 790 amino acids comprising the hemoglobin receptor gene. In this embodiment of the invention, the nucleotide sequence of the *N. meningitidis* hemoglobin receptor gene is the sequence depicted in Figure 7 (SEQ ID No.:3). It will be understood that the *N. meningitidis* gene as disclosed herein is defined, insofar as is necessary, by the amino acid sequence of the protein encoded therein,

said amino acid sequence being represented in Figure 7 (SEQ. ID No.:4). Thus, it will be understood that the particular nucleotide sequence depicted in Figure 7 (SEQ. ID. No.:3) is but one of a number of equivalent nucleotide sequences that encode the hemoglobin receptor protein, due to the degeneracy of the genetic code, and that all such alternative, equivalent nucleotide sequences are hereby explicitly encompassed within the disclosed nucleotide sequences of the invention. Also included herein are any mutant or allelic variations of this nucleotide sequence, either naturally occurring or the product of *in vitro* chemical or genetic modification. Each such variant will be understood to have essentially the same nucleotide sequence as the nucleotide sequence of the corresponding *N. meningitidis* hemoglobin receptor protein disclosed herein.

In another particularly preferred embodiment of this aspect of the invention, the hemoglobin receptor protein gene is isolated from *Neisseria meningitidis*, serotype B. In a particular example of this embodiment, the nucleic acid comprises a 2376 basepair (bp) polymerase chain reaction-amplified fragment of *N. meningitidis*, serotype A genomic DNA. In this embodiment, the nucleotide sequence comprises an open reading frame of 2373 nucleotides of *N. meningitidis* genomic DNA encoding 791 amino acids comprising the hemoglobin receptor gene. In this embodiment of the invention, the nucleotide sequence of the *N. meningitidis* hemoglobin receptor gene is the sequence depicted in Figure 8 (SEQ ID No:5). It will be understood that the *N. meningitidis* gene as disclosed herein is defined, insofar as is necessary, by the amino acid sequence of the protein encoded therein, said amino acid sequence being represented in Figure 8 (SEQ. ID No.:6). Thus, it will be understood that the particular nucleotide sequence depicted in Figure 8 (SEQ. ID. No.:5) is but one of a number of equivalent nucleotide sequences that encode the hemoglobin receptor protein, due to the degeneracy of the genetic code, and that all such alternative, equivalent nucleotide sequences are hereby explicitly encompassed within the disclosed nucleotide sequences of the invention. Also included herein are any mutant or allelic variations of this nucleotide sequence, either naturally occurring or the product of *in vitro* chemical or genetic modification. Each such variant will be understood to have essentially the same nucleotide sequence as the nucleotide

sequence of the corresponding *N. meningitidis* hemoglobin receptor protein disclosed herein.

In yet other preferred embodiments, the invention provides nucleic acid encoding a hemoglobin receptor protein gene isolated from *Neisseria gonorrhoeae*.  
5 In a particular example of this embodiment, the nucleic acid comprises a 2378 basepair (bp) polymerase chain reaction-amplified fragment of *N. gonorrhoeae* genomic DNA. In this embodiment, the nucleotide sequence comprises an open reading frame of 2373 nucleotides of *N. gonorrhoeae* genomic DNA encoding 791 amino acids comprising the hemoglobin receptor gene. In this embodiment of the  
10 invention, the nucleotide sequence of the *N. gonorrhoeae* hemoglobin receptor gene is the sequence depicted in Figure 9 (SEQ ID No:7). It will be understood that the *N. gonorrhoeae* gene as disclosed herein is defined, insofar as is necessary, by the amino acid sequence of the protein encoded therein, said amino acid sequence being represented in Figure 9 (SEQ. ID No.:8). Thus, it will be understood that the  
15 particular nucleotide sequence depicted in Figure 9 (SEQ. ID. No.:7) is but one of a number of equivalent nucleotide sequences that encode the hemoglobin receptor protein, due to the degeneracy of the genetic code, and that all such alternative, equivalent nucleotide sequences are hereby explicitly encompassed within the disclosed nucleotide sequences of the invention. Also included herein are any mutant  
20 or allelic variations of this nucleotide sequence, either naturally occurring or the product of *in vitro* chemical or genetic modification. Each such variant will be understood to have essentially the same nucleotide sequence as the nucleotide sequence of the corresponding *N. gonorrhoeae* hemoglobin receptor protein disclosed herein.

25 The invention also provides bacterial hemoglobin receptor proteins. In a preferred embodiment, the bacterial hemoglobin receptor protein is isolated from bacteria of *Neisseria* species. In a particularly preferred embodiment, the hemoglobin receptor protein is isolated from *Neisseria meningitidis*. In a particular example of this embodiment, the protein is derived from *N. meningitidis*, serotype  
30 C and comprises an amino acid sequence of 792 amino acids. In this embodiment of the invention, the amino acid sequence of the *N. meningitidis*, serotype C hemoglobin receptor protein is the sequence depicted in Figure 2 (SEQ ID No:2).

In another example of this embodiment, the protein is derived from *N. meningitidis*, serotype A and comprises an amino acid sequence of 790 amino acids. In this embodiment of the invention, the amino acid sequence of the *N. meningitidis*, serotype A hemoglobin receptor protein is the sequence depicted in Figure 7 (SEQ ID No:4). In yet another example of this embodiment, the protein is derived from *N. meningitidis*, serotype B and comprises an amino acid sequence of 791 amino acids. In this embodiment of the invention, the amino acid sequence of the *N. meningitidis*, serotype B hemoglobin receptor protein is the sequence depicted in Figure 8 (SEQ ID No:6). The invention also provides hemoglobin receptor protein derived from *N. gonorrhoeae*. In this embodiment of the invention, the protein comprises an amino acid sequence of 791 amino acids, and the amino acid sequence of the *N. gonorrhoeae* hemoglobin receptor protein is the sequence depicted in Figure 9 (SEQ ID No:8). Also explicitly encompassed within the scope of this invention are related bacterial hemoglobin receptor proteins, particularly such proteins isolated from *Neisseria* species, having essentially the same amino acid sequence and substantially the same biological properties as the hemoglobin receptor protein encoded by the *N. meningitidis* and *N. gonorrhoeae* nucleotide sequences described herein.

In another aspect, the invention provides a homogeneous preparation of an approximately 85.5 kiloDalton (kD) bacterial hemoglobin receptor protein or derivative thereof, said size being understood to be the size of the protein before any post-translational modifications thereof. Also provided is a 90kD embodiment of the receptor as determined by sodium dodecyl sulfate/ polyacrylamide gel electrophoresis under reducing conditions. In a preferred embodiment, the bacterial hemoglobin receptor protein is isolated from bacteria of *Neisseria* species. In a particularly preferred embodiment, the hemoglobin receptor protein is isolated from *Neisseria meningitidis*. In one embodiment of this aspect of the invention, the protein is isolated from *N. meningitidis*, serotype C and the amino acid sequence of the bacterial hemoglobin receptor protein or derivative thereof preferably is the amino acid sequence of the hemoglobin receptor protein shown in Figure 2 (SEQ ID No:2). In a second embodiment of this aspect of the invention, the protein is isolated from *N. meningitidis*, serotype A and the amino acid sequence of the bacterial hemoglobin

receptor protein or derivative thereof preferably is the amino acid sequence of the hemoglobin receptor protein shown in Figure 7 (SEQ ID No:4). In a third embodiment of this aspect of the invention, the protein is isolated from *N. meningitidis*, serotype B and the amino acid sequence of the bacterial hemoglobin receptor protein or derivative thereof preferably is the amino acid sequence of the hemoglobin receptor protein shown in Figure 8 (SEQ ID No:6). The invention also provides a homogeneous preparation of a bacterial hemoglobin receptor protein isolated from *N. gonorrhoeae*. In a preferred embodiment, the amino acid sequence of the bacterial hemoglobin receptor protein or derivative thereof preferably is the amino acid sequence of the hemoglobin receptor protein shown in Figure 9 (SEQ ID No:8).

This invention provides nucleotide probes derived from the nucleotide sequences herein provided. The invention includes probes isolated from either complementary DNA (cDNA) copies of bacterial messenger RNA (mRNA) or bacterial genomic DNA (gDNA), as well as probes made synthetically or by *in vitro* amplification methods using the sequence information provided herein. The invention specifically includes but is not limited to oligonucleotide, nick-translated, random primed, or *in vitro* amplified probes made using cDNA or genomic clones embodying the invention, and oligonucleotide and other synthetic probes synthesized chemically using the nucleotide sequence information of cDNA or genomic clone embodiments of the invention.

It is a further object of this invention to provide such nucleic acid hybridization probes to detect the presence of bacteria of *Neisseria* species, particularly *N. meningitidis* and *N. gonorrhoeae*, in a biological sample in the diagnosis of a *Neisseria* infection in a human. Such a biological sample preferably includes blood, urine, semen, mucus, cerebrospinal fluid, peritoneal fluid and ascites fluids, as well as cell scrapings from the epithelium of the mouth, urethra, anus and rectum, and other organs.

The present invention also includes peptides encoded by the nucleotide sequences comprising the nucleic acid embodiments of the invention. The invention includes either naturally occurring or synthetic peptides which may be used as antigens for the production of hemoglobin receptor protein-specific antibodies. The



invention also comprises such antibodies, preferably monoclonal antibodies, and cells and cultures of cells producing such antibodies.

Thus, the invention also provides antibodies against and epitopes of bacterial hemoglobin receptor proteins of the invention. It is an object of the present invention to provide antibodies that are immunologically reactive to the bacterial hemoglobin receptor proteins of the invention. It is a particular object to provide monoclonal antibodies against these bacterial hemoglobin receptor proteins. In a preferred embodiment, antibodies provided are raised against bacterial hemoglobin receptor protein isolated from bacteria of *Neisseria* species. In a particularly preferred embodiment, such antibodies are specific for the hemoglobin receptor protein isolated from *Neisseria meningitidis* serotypes A, B or C. In additional particularly preferred embodiment, such antibodies are specific for the hemoglobin receptor protein isolated from *Neisseria gonorrhoeae*.

Hybridoma cell lines producing such antibodies are also objects of the invention. It is envisioned at such hybridoma cell lines may be produced as the result of fusion between a non-immunoglobulin producing mouse myeloma cell line and spleen cells derived from a mouse immunized with purified hemoglobin receptor protein or a cell expressing antigens or epitopes of bacterial hemoglobin receptor proteins of the invention. The present invention also provides hybridoma cell lines that produce such antibodies, and can be injected into a living mouse to provide an ascites fluid from the mouse that is comprised of such antibodies. In a preferred embodiment, antibodies provided are raised against bacterial hemoglobin receptor protein isolated from bacteria of *Neisseria* species. In a particularly preferred embodiment, such antibodies are specific for the hemoglobin receptor protein isolated from *Neisseria meningitidis*, serotypes A, B or C. In additional particularly preferred embodiment, such antibodies are specific for the hemoglobin receptor protein isolated from *Neisseria gonorrhoeae*.

It is a further object of the invention to provide immunologically-active epitopes of the bacterial hemoglobin receptor proteins of the invention. Chimeric antibodies immunologically reactive against the bacterial hemoglobin receptor proteins of the invention are also within the scope of this invention. In a preferred embodiment, antibodies and epitopes provided are raised against or derived from

bacterial hemoglobin receptor protein isolated from bacteria of *Neisseria* species. In a particularly preferred embodiment, such antibodies and epitopes are specific for the hemoglobin receptor protein isolated from *Neisseria meningitidis*, serotypes A, B or C. In additional particularly preferred embodiment, such antibodies and  
5 epitopes are specific for the hemoglobin receptor protein isolated from *Neisseria gonorrhoeae*.

The present invention provides recombinant expression constructs comprising a nucleic acid encoding a bacterial hemoglobin receptor protein wherein the construct is capable of expressing the encoded hemoglobin receptor protein in cultures of cells  
10 transformed with the construct. Preferred embodiments of such constructs comprise the *N. meningitidis*, serotype C hemoglobin receptor gene depicted in Figure 2 (SEQ ID No.:1), such constructs being capable of expressing the bacterial hemoglobin receptor protein encoded therein in cells transformed with the construct. Additional preferred embodiments of such constructs comprise the *N. meningitidis*, serotype A  
15 hemoglobin receptor gene depicted in Figure 7 (SEQ ID No.:3), such constructs being capable of expressing the bacterial hemoglobin receptor protein encoded therein in cells transformed with the construct. Further additional preferred embodiments of such constructs comprise the *N. meningitidis*, serotype B hemoglobin receptor gene depicted in Figure 8 (SEQ ID No.:5), such constructs being capable  
20 of expressing the bacterial hemoglobin receptor protein encoded therein in cells transformed with the construct. The invention also provides recombinant expression constructs encoding a hemoglobin receptor protein gene isolated from *N. gonorrhoeae*. In a particularly preferred embodiment, such constructs comprise the *N. gonorrhoeae* hemoglobin receptor gene depicted in Figure 9 (SEQ ID No.:7), the  
25 constructs being capable of expressing the bacterial hemoglobin receptor protein encoded therein in cells transformed with the construct.

The invention also provides cultures of cells, preferably bacterial cells, having been transformed with the recombinant expression constructs of the invention, each such cultures being capable of and in fact expressing the bacterial hemoglobin  
30 receptor protein encoded in the transforming construct.

The present invention also includes within its scope protein preparations of prokaryotic cell membranes containing the bacterial hemoglobin receptor protein of

the invention, derived from cultures of prokaryotic cells transformed with the recombinant expression constructs of the invention.

5 The invention also provides diagnostic reagents and methods for using such reagents for detecting the existence of an infection in a human, with bacteria of a *Neisseria* species. In preferred embodiments, such diagnostic reagents comprise antibodies that are immunologically reactive with a bacterial hemoglobin receptor protein. In a preferred embodiment, such antibodies are raised against a bacterial hemoglobin receptor protein isolated from bacteria of *Neisseria* species. In a particularly preferred embodiment, such antibodies are specific for the hemoglobin receptor protein isolated from *Neisseria meningitidis*, serotypes A, B or C. In  
10 additional particularly preferred embodiments, such antibodies are specific for the hemoglobin receptor protein isolated from *Neisseria gonorrhoeae*.

In yet another embodiment of this aspect of the invention are provided diagnostic reagents and methods for using such reagents wherein said reagents are  
15 nucleic acid hybridization probes comprising a bacterial hemoglobin receptor gene. In a preferred embodiment, the bacterial hemoglobin receptor protein gene is isolated from bacteria of *Neisseria* species. In a particularly preferred embodiment, the hemoglobin receptor protein gene is isolated from *Neisseria meningitidis*. In particular examples of this embodiment of the invention, the nucleic acid probes  
20 comprise a specifically-hybridizing fragment of a 3.3 kilobase (kb) *Bam*HI/*Hind*III fragment of *N. meningitidis*, serotype C genomic DNA. In this embodiment, the nucleotide sequence comprises all or a specifically-hybridizing fragment of an open reading frame of 2376 nucleotides of *N. meningitidis*, serotype C genomic DNA encoding 792 amino acids comprising the hemoglobin receptor gene. In this  
25 embodiment of the invention, the nucleotide sequence of the *N. meningitidis*, serotype C hemoglobin receptor gene is the sequence depicted in Figure 2 (SEQ ID No:1). In another example of this embodiment of the invention, the nucleic acid probes comprise a specifically-hybridizing fragment of a 2373bp, polymerase chain reaction-amplified fragment of *N. meningitidis*, serotype A genomic DNA. In this  
30 embodiment, the nucleotide sequence comprises all or a specifically-hybridizing fragment of an open reading frame of 2370 nucleotides of *N. meningitidis*, serotype A genomic DNA encoding 790 amino acids comprising the hemoglobin receptor

gene. In this embodiment of the invention, the nucleotide sequence of the *N. meningitidis*, serotype A hemoglobin receptor gene is the sequence depicted in Figure 7 (SEQ ID No:3). In yet another example of this embodiment of the invention, the nucleic acid probes comprise a specifically-hybridizing fragment of a 2376bp, polymerase chain reaction-amplified fragment of *N. meningitidis*, serotype B genomic DNA. In this embodiment, the nucleotide sequence comprises all or a specifically-hybridizing fragment of an open reading frame of 2373 nucleotides of *N. meningitidis*, serotype B genomic DNA encoding 791 amino acids comprising the hemoglobin receptor gene. In this embodiment of the invention, the nucleotide sequence of the *N. meningitidis*, serotype B hemoglobin receptor gene is the sequence depicted in Figure 8 (SEQ ID No:5). The invention also provides nucleic acid hybridization probes comprising a bacterial hemoglobin receptor gene isolated from *N. gonorrhoeae*. In a preferred embodiment of this aspect of the invention, the nucleic acid probes comprise a specifically-hybridizing fragment of a 2378bp, polymerase chain reaction-amplified fragment of *N. gonorrhoeae* genomic DNA. In this embodiment, the nucleotide sequence comprises all or a specifically-hybridizing fragment of an open reading frame of 2373 nucleotides of *N. gonorrhoeae* genomic DNA encoding 791 amino acids comprising the hemoglobin receptor gene. In this embodiment of the invention, the nucleotide sequence of the *N. gonorrhoeae* hemoglobin receptor gene is the sequence depicted in Figure 9 (SEQ ID No:7). It will be understood that the term "specifically-hybridizing" when used to describe a fragment of a nucleic acid encoding a bacterial hemoglobin receptor gene is intended to mean that nucleic acid hybridization of such a fragment is stable under high stringency conditions of hybridization and washing as the term "high stringency" would be understood by those having skill in the molecular biological arts.

Also provided by the invention are therapeutic agents and methods for using such agents for treating the an infection in a human, with bacteria of a *Neisseria* species. In preferred embodiments, such agents comprise antibodies that are immunologically reactive with a bacterial hemoglobin receptor protein. In a preferred embodiment, such antibodies are raised against a bacterial hemoglobin receptor protein isolated from bacteria of *Neisseria* species. In a particularly preferred embodiment, such antibodies are specific for the hemoglobin receptor

protein isolated from *Neisseria meningitidis*, serotypes A, B or C. In additional preferred embodiments, such antibodies are specific for the hemoglobin receptor protein isolated from *Neisseria gonorrhoeae*. Therapeutic agents provided in this aspect of the invention comprise such antibodies in a pharmaceutically-acceptable carrier, along with appropriate adjuvants and the like. In additional embodiments, such antibodies are covalently conjugated to a bactericidal or bacteriostatic agent effective against bacteria of *Neisseria* species, preferably *N. meningitidis* and *N. gonorrhoeae*.

In yet another embodiment of this aspect of the invention are provided therapeutic reagents and methods for using such reagents wherein said reagents comprise recombinant expression constructs of the invention, or a homologue thereof that expresses the nucleic acid encoding a hemoglobin receptor in an antisense orientation. In a preferred embodiment, the bacterial hemoglobin receptor protein gene is isolated from bacteria of *Neisseria* species. In a particularly preferred embodiment, the hemoglobin receptor protein gene is isolated from *Neisseria meningitidis*. In particular examples of this embodiment of the invention, the nucleic acids comprise a specifically-hybridizing fragment of a 3.3 kilobase (kb) *Bam*HI/*Hind*III fragment of *N. meningitidis*, serotype C genomic DNA. In this embodiment, the nucleotide sequence comprises all or a specifically-hybridizing fragment of an open reading frame of 2376 nucleotides of *N. meningitidis*, serotype C genomic DNA encoding 792 amino acids comprising the hemoglobin receptor gene. In this embodiment of the invention, the nucleotide sequence of the *N. meningitidis*, serotype C hemoglobin receptor gene is the sequence depicted in Figure 2 (SEQ ID No:1). In another example of this embodiment of the invention, the nucleic acid probes comprise a specifically-hybridizing fragment of a 2373bp, polymerase chain reaction-amplified fragment of *N. meningitidis*, serotype A genomic DNA. In this embodiment, the nucleotide sequence comprises all or a specifically-hybridizing fragment of an open reading frame of 2370 nucleotides of *N. meningitidis*, serotype A genomic DNA encoding 790 amino acids comprising the hemoglobin receptor gene. In this embodiment of the invention, the nucleotide sequence of the *N. meningitidis*, serotype A hemoglobin receptor gene is the sequence depicted in Figure 7 (SEQ ID No:3). In yet another example of this

embodiment of the invention, the nucleic acid probes comprise a specifically-hybridizing fragment of a 2376bp, polymerase chain reaction-amplified fragment of *N. meningitidis*, serotype B genomic DNA. In this embodiment, the nucleotide sequence comprises all or a specifically-hybridizing fragment of an open reading frame of 2373 nucleotides of *N. meningitidis*, serotype B genomic DNA encoding 791 amino acids comprising the hemoglobin receptor gene. In this embodiment of the invention, the nucleotide sequence of the *N. meningitidis*, serotype B hemoglobin receptor gene is the sequence depicted in Figure 8 (SEQ ID No:5). The invention also provides recombinant expression constructs of the invention, or a homologue thereof that expresses the nucleic acid encoding a hemoglobin receptor in an antisense orientation, wherein the nucleic acid encodes a bacterial hemoglobin receptor gene isolated from *N. gonorrhoeae*. In a preferred embodiment of this aspect of the invention, the nucleic acid probes comprise a specifically-hybridizing fragment of a 2378bp, polymerase chain reaction-amplified fragment of *N. gonorrhoeae* genomic DNA. In this embodiment, the nucleotide sequence comprises all or a specifically-hybridizing fragment of an open reading frame of 2373 nucleotides of *N. gonorrhoeae* genomic DNA encoding 791 amino acids comprising the hemoglobin receptor gene. In this embodiment of the invention, the nucleotide sequence of the *N. gonorrhoeae* hemoglobin receptor gene is the sequence depicted in Figure 9 (SEQ ID No:7).

The invention also provides a method for screening compounds for their ability to inhibit, facilitate or modulate the biochemical activity of a bacterial hemoglobin receptor protein of the invention, for use in the *in vitro* screening of novel agonist and antagonist compounds and novel bactericidal and bacteriostatic agents specific for the hemoglobin receptor protein. In preferred embodiments, cells transformed with a recombinant expression construct of the invention are contacted with such a compound, and the binding capacity of the compounds, as well as the effect of the compound on binding of other, known hemoglobin receptor agonists such as hemoglobin and hemin, and antagonists, is assayed. Additional preferred embodiments comprise quantitative analyses of such effects.

The present invention is also useful for the detection of bactericidal and/or bacteriostatic analogues, agonists or antagonists, known or unknown, of a bacterial

hemoglobin receptor protein, preferably derived from bacteria of *Neisseria* species, most preferably isolated from *N. meningitidis*, wherein such compounds are either naturally occurring or embodied as a drug.

5 The invention also provides vaccines for immunizing a human against infection with pathogenic bacteria of *Neisseria* species, the vaccines comprising the hemoglobin binding proteins of the invention or antigenic fragments thereof. In a preferred embodiment, the vaccines of the invention comprise cells expressing a hemoglobin receptor binding protein of the invention, or an antigenic fragment thereof, preferably wherein said cells are attenuated varieties of cells adapted for  
10 growth in humans, *i.e.*, wherein such cells are non-pathogenic and do not cause bacteremia, endotoxemia or sepsis. Examples of such attenuated varieties of cells include attenuated strains of *Salmonella* species, for example *Salmonella typhi* and *Salmonella typhimurium*, as well as other attenuated bacterial species. Also provided by the invention are recombinant expression constructs as disclosed herein useful *per*  
15 *se* as vaccines, for introduction into an animal and production of an immunologic response to bacterial hemoglobin receptor protein antigens encoded therein.

Specific preferred embodiments of the present invention will become evident from the following more detailed description of certain preferred embodiments and the claims.

20

### DESCRIPTION OF THE DRAWINGS

The foregoing and other objects of the present invention, the various features thereof, as well as the invention itself may be more fully understood from the following description, when read together with the accompanying drawings in which:

25 Figure 1 is a schematic drawing of the restriction enzyme digestion map of a *N. meningitidis* cosmid clone and subclones thereof derived as described in Example 2.

Figure 2 illustrates the nucleotide (SEQ ID No.:1) and deduced amino acid (SEQ ID No.:2) sequences of the *N. meningitidis* hemoglobin receptor protein  
30 encoded in a 3.3 kb *Bam*HI/*Hind*III DNA fragment.

Figure 3 presents a photograph of a stained SDS/ 10% PAGE electrophoresis gel showing the results of *in vitro* expression of the *N. meningitidis* hemoglobin

receptor gene product as an approximately 90 kilodalton protein, and  $\beta$ -lactamase protein having a molecular weight of about 30.0 kilodaltons used as a molecular weight marker.

5 Figure 4 presents an amino acid sequence comparison between portions of the *N. meningitidis* transferrin receptor Tbp1 (SEQ ID No.:9), the *N. meningitidis* lactoferrin receptor LbpA (SEQ ID No.:10), and *N. meningitidis* hemoglobin receptor HmbR (SEQ ID No.:2).

10 Figure 5 illustrates Southern hybridization analysis of chromosomal DNA from *N. meningitidis* 8013 and the MC8013*hmbR* mutant using a *Bam*HI-*Sal*I fragment of the *hmb* gene as probe labeled using a DIG nonradioactive DNA labelling and detection kit (Boehringer Mannheim Biochemicals, Indianapolis, IN). Lane 1 contains DNA from *N. meningitidis* strain MC8013, digested with *Cl*aI; lane 2 is MC8013*hmbR* DNA digested with *Cl*aI; lane 3, is MC8013 DNA digested with *Bam*HI and *Sal*I; and lane 4 is MC8013*hmbR* DNA digested with *Bam*HI and *Sal*I.

15 Figure 6 is a graph describing the course of infection using *N. meningitidis* wild type (MC8013) and *hmbR* mutant strains in an *in vivo* rat infant infection model. Each strain was injected intraperitoneally ( $2 \times 10^6$  CFU) into three infant inbred Lewis rats. The results represent the average of two similarly-performed experiments.

20 Figure 7 illustrates the nucleotide (SEQ ID No.:3) and deduced amino acid (SEQ ID No.:4) sequences of the *N. meningitidis*, serotype A hemoglobin receptor protein encoded on a 2373bp polymerase chain reaction-amplified DNA fragment.

25 Figure 8 illustrates the nucleotide (SEQ ID No.:5) and deduced amino acid (SEQ ID No.:6) sequences of the *N. meningitidis*, serotype B hemoglobin receptor protein encoded on a 2376bp polymerase chain reaction-amplified DNA fragment.

Figure 9 illustrates the nucleotide (SEQ ID No.:7) and deduced amino acid (SEQ ID No.:8) sequences of the *N. gonorrhoeae* hemoglobin receptor protein encoded on a 2376bp polymerase chain reaction-amplified DNA fragment.

30 Figure 10 represents a schematic of a nucleic acid sequence comparison between the hemoglobin receptor proteins derived from *N. meningitidis*, serotypes A (SEQ ID No.:3), B (SEQ ID No.:5) and C (SEQ ID No.:1) and from *N. gonorrhoeae* (SEQ ID No.:7), wherein the direction of transcription of the genes is



in the direction of the arrow, and the following abbreviations refer to restriction endonuclease sites: H represents *Hind*III; N represents *Not*I; Bg represents *Bgl*II; Bs represents *Bss*HI; Nr represents *Nru*I; Cl represents *Cl*aI; P represents *Pst*I; Sa represents *Sac*I; Av represents *Ava*I; B represents *Bam*HI; S represents *Sa*II; EV represents *Eco*RV; Sh represents *Sph*I; and Sy represents *Syl*.

Figure 11 presents an amino acid sequence comparison between the hemoglobin receptor proteins derived from *N. meningitidis*, serotypes A (SEQ ID No.:4), B (SEQ ID No.:6) and C (SEQ ID No.:2) and from *N. gonorrhoeae* (SEQ ID No.:8).

### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The term "bacterial hemoglobin receptor" as used herein refers to bacterial proteins comprising the outer membrane of Gram negative bacteria, which specifically mediate transit of hemoglobin-derived heme, as well as heme from other sources, through the outer membrane of such bacteria and into the periplasmic space. The bacterial hemoglobin receptor proteins of the invention are characterized by, first, an amino acid sequence that is essentially the sequence depicted in Figures 2 (SEQ ID No.:2), 7 (SEQ ID No.:4), 8 (SEQ ID No.:6) and 9 (SEQ ID No.:8). The bacterial hemoglobin receptor proteins of the invention are further characterized by having substantially the same biological activity as a protein having the amino acid sequence depicted in Figures 2 (SEQ ID No.:2), 7 (SEQ ID No.:4), 8 (SEQ ID No.:6) and 9 (SEQ ID No.:8). This definition is intended to encompass naturally-occurring variants and mutant proteins, as well as genetically engineered variants made by man.

Cloned, isolated and purified nucleic acid provided by the present invention may encode a bacterial hemoglobin receptor protein of any *Neisseria* species of origin, including, most preferably, *Neisseria meningitidis* species and serotypes thereof and *Neisseria gonorrhoeae* species.

The nucleic acid hybridization probes provided by the invention comprise DNA or RNA having all or a specifically-hybridizing fragment of the nucleotide sequence of the hemoglobin receptor protein as depicted in Figures 2 (SEQ ID No.:1), 7 (SEQ ID No.:3), 8 (SEQ ID No.:5) and 9 (SEQ ID No.:7), or any portion

thereof effective in nucleic acid hybridization. Mixtures of such nucleic acid hybridization probes are also within the scope of this embodiment of the invention. Nucleic acid probes as provided herein are useful for detecting the presence of a bacteria, *inter alia*, in a human as the result of an infection, in contaminated biological samples and specimens, in foodstuffs and water supplies, or in any substance that may come in to contact with the human. Specific hybridization will be understood to mean that the nucleic acid probes of the invention are capable of forming stable, specific hybridization to bacterially-derived DNA or RNA under conditions of high stringency, as the term "high stringency" would be understood by those with skill in the art (*see, for example*, Sambrook *et al.*, 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. and Hames and Higgins, eds., 1985, Nucleic Acid Hybridization, IRL Press, Oxford, U.K.). Hybridization will be understood to be accomplished using well-established techniques, including but not limited to Southern blot hybridization, Northern blot hybridization, *in situ* hybridization and Southern hybridization to polymerase chain reaction product DNAs. The invention will thus be understood to provide oligonucleotides, specifically, pairs of oligonucleotides, for use as primers in support of *in vitro* amplification of bacterial hemoglobin receptor genes and mRNA transcripts.

The production of proteins such as bacterial hemoglobin receptor proteins from cloned genes by genetic engineering means is well known in this art. The discussion which follows is accordingly intended as an overview of this field, and is not intended to reflect the full state of the art. It will be understood from the following discussion that the hemoglobin receptor protein genes of this invention are particularly advantageous, since expression of such proteins by bacteria, including non-*Neisseria* species of bacteria, can complement certain auxotrophic mutants of said transformed bacteria otherwise unable to subsist absent supplementation of the growth media with iron (III).

DNA encoding a bacterial hemoglobin receptor protein, in view of the instant disclosure, by chemical synthesis, by screening reverse transcripts of mRNA from appropriate cells, by screening genomic libraries from appropriate cells, or by combinations of these procedures, as illustrated below. Screening of mRNA or

genomic DNA may be carried out with oligonucleotide probes generated from the nucleic acid sequence information from the bacterial hemoglobin receptor protein disclosed herein. Probes may be labeled with a detectable group such as a fluorescent group, a radioactive atom or a chemiluminescent group in accordance with know procedures and used in conventional hybridization assays, as described in greater detail in the Examples below. In the alternative, bacterial hemoglobin receptor protein-encoding nucleic acids may be obtained by use of the polymerase chain reaction (PCR) procedure, using appropriate pairs of PCR oligonucleotide primers corresponding to nucleic acid sequence information derived from a bacterial hemoglobin receptor protein as provided herein. See U.S. Patent Nos. 4,683,195 to Mullis *et al.* and 4,683,202 to Mullis, as specifically disclosed herein in Example 9 below. In another alternative, such bacterial hemoglobin receptor protein-encoding nucleic acids may be isolated from auxotrophic cells transformed with a bacterial hemoglobin receptor protein gene, thereby relieved of the nutritional requirement for uncomplexed iron (III).

Any bacterial hemoglobin receptor protein of the invention may be synthesized in host cells transformed with a recombinant expression construct comprising a nucleic acid encoding the bacterial hemoglobin receptor protein. Such recombinant expression constructs can also be comprised of a vector that is a replicable DNA construct. Vectors are used herein either to amplify DNA encoding a bacterial hemoglobin receptor protein and/or to express DNA encoding a bacterial hemoglobin receptor protein. For the purposes of this invention, a recombinant expression construct is a replicable DNA construct in which a nucleic acid encoding a bacterial hemoglobin receptor protein is operably linked to suitable control sequences capable of effecting the expression of the bacterial hemoglobin receptor protein in a suitable host cell.

The need for such control sequences will vary depending upon the host cell selected and the transformation method chosen. Generally, bacterial control sequences include a transcriptional promoter, an optional operator sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding sites (the Shine-Delgarno sequence), and sequences which control the termination of transcription and translation. Amplification vectors do not require expression control

domains. All that is needed is the ability to replicate in a host, usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants. *See, Sambrook et al., 1989, ibid.*

5 Vectors useful for practicing the present invention include plasmids and virus-derived constructs, including phage and particularly bacteriophage, and integratable DNA fragments (i.e., fragments integratable into the host genome by homologous recombination). The vector replicates and functions independently of the host genome, or may, in some instances, integrate into the genome itself. Suitable vectors will contain replicon and control sequences which are derived from species  
10 compatible with the intended expression host. A preferred vector is pLAFR2 (*see Riboli et al., 1991, Microb. Pathogen. 10: 393-403*).

Transformed host cells are cells which have been transformed or transfected with recombinant expression constructs made using recombinant DNA techniques and comprising nucleic acid encoding a bacterial hemoglobin receptor protein. Preferred  
15 host cells are cells of *Neisseria* species, particularly *N. meningitidis*, as well as *Salmonella typhi* and *Salmonella typhimurium* species, and *Escherichia coli* auxotrophic mutant cells (*hemA aroB*). Transformed host cells may express the bacterial hemoglobin receptor protein, but host cells transformed for purposes of cloning or amplifying nucleic acid hybridization probe DNA need not express the  
20 receptor protein. When expressed, the bacterial hemoglobin receptor protein of the invention will typically be located in the host cell outer membrane. *See, Sambrook et al., ibid.*

Cultures of bacterial cells, particularly cells of *Neisseria* species, and certain *E. coli* mutants, are a desirable host for recombinant bacterial hemoglobin receptor  
25 protein synthesis. In principal, any bacterial cell auxotrophic for uncomplexed iron (III) is useful for selectively growing bacterial hemoglobin receptor protein-transformed cells. However, for this purpose, well-characterized auxotrophs, such as *E. coli hemA aroB* mutants are preferred.

The invention provides homogeneous compositions of a bacterial hemoglobin  
30 receptor protein produced by transformed cells as provided herein. Each such homogeneous composition is intended to be comprised of a bacterial hemoglobin receptor protein that comprises at least 90% of the protein in such a homogenous

composition. The invention also provides membrane preparations from cells expressing a bacterial hemoglobin receptor protein as the result of transformation with a recombinant expression construct of the invention, as described herein.

5 Bacterial hemoglobin receptor proteins, peptide fragments thereof and membranes derived from cells expressing such proteins in accordance with the present invention may be used for the production of vaccines effective against bacterial infections in a human, with pathogenic microorganisms expressing such bacterial hemoglobin receptor proteins. Such vaccines preferably would be effective in raising an immunological response against bacteria of *Neisseria* species, most  
10 preferably *N. meningitidis* and *N. gonorrhoeae*. Also encompassed within the vaccines provided by the invention are recombinant expression constructs as disclosed herein useful *per se* as vaccines, for introduction into an animal and production of an immunologic response to bacterial hemoglobin receptor protein antigens encoded therein.

15 Preparation of vaccines which contain polypeptide or polynucleotide sequences as active ingredients is well understood in the art. Typically, such vaccines are prepared as injectables, either as liquid solutions or suspensions. However, solid forms suitable for solution in, or suspension in, liquid prior to injection may also be prepared. The preparation may also be emulsified. The active  
20 immunogenic ingredient is often mixed with excipients which are pharmaceutically acceptable and compatible with the active ingredient. Suitable excipients are, for example, water, saline, dextrose, glycerol, ethanol, or the like and combinations thereof. In addition, if desired, the vaccine may contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, or adjuvants  
25 which enhance the effectiveness of the vaccine. The vaccines are conventionally administered parenterally, by injection, for example, either subcutaneously or intramuscularly. Additional formulations which are suitable for other modes of administration include suppositories and, in some cases, oral formulations. For suppositories, traditional binders and carriers may include, for example, polyalkylene  
30 glycols or triglycerides; such suppositories may be formed from mixtures containing the active ingredient in the range of 0.5% to 10%, preferably 1 to 2%. Oral formulations include such normally employed excipients as, for example,

pharmaceutical grades of manitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate and the like. These compositions take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and contain 10% to 95% of active ingredient, preferably 25 to 70%.

The polypeptides of the invention may be formulated into the vaccine as neutral or salt forms. Pharmaceutically acceptable salts, include the acid additional salts (formed with the free amino groups of the peptide) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups may also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine, and the like.

In another embodiment, such vaccines are provided wherein the bacterial hemoglobin receptor proteins or peptide fragments thereof are present in the intact cell membranes of cells expressing such proteins in accordance with the present invention. In preferred embodiments, cells useful in these embodiments include attenuated varieties of cells adapted to growth in humans. Most preferably, said cells are attenuated varieties of cells adapted for growth in humans, *i.e.*, wherein such cells do not cause frank disease or other pathological conditions, such as bacteremia, endotoxemia or sepsis. For the purposes of this invention, "attenuated" cells will be understood to encompass prokaryotic and eukaryotic cells that do not cause infection, disease, septicemia, endotoxic shock, pyrogenic shock, or other serious and adverse reactions to administration of vaccines to an animal, most preferably a human, when such cells are introduced into the animal, whether such cells are viable, living, heat-, chemically- or genetically attenuated or inactivated, or dead. It will be appreciated by those with skill in this art that certain minor side-effects of vaccination, such as short-term fever, muscle discomfort, general malaise, and other well-known reactions to vaccination using a variety of different types of vaccines, can be anticipated as accompanying vaccination of an animal, preferably a human, using the vaccines of the invention. Such acute, short-term and non-life-threatening side effects are

encompassed in the instant definition of the vaccines of the invention, and vaccines causing such side-effects fall within the definition of "attenuated" presented herein. Preferred examples of such attenuated cells include attenuated varieties of *Salmonella* species, preferably *Salmonella typhi* and *Salmonella typhimurium*, as well as other attenuated bacterial species. It will be specifically understood that these embodiments of the vaccines of the invention encompass so-called "live" attenuated cell preparations as well as heat- or chemically-inactivated cell preparations.

In other embodiments of the invention are provided vaccines that are DNA vaccines, comprising the nucleic acids of the invention in recombinant expression constructs competent to direct expression of hemoglobin receptor proteins when introduced into an animal. In preferred embodiments, such DNA vaccines comprise recombinant expression constructs wherein the hemoglobin receptor-encoding nucleic acids of the invention are operably linked to promoter elements, most preferably the early gene promoter of cytomegalovirus or the early gene promoter of simian virus 40. DNA vaccines of the invention are preferably administered by intramuscular injection, but any appropriate route of administration, including oral, transdermal, rectal, nasal, aerosol administration into lung, or any other clinically-acceptable route of administration can be used by those with skill in the art.

In general, the vaccines of the invention are administered in a manner compatible with the dosage formulation, and in such amount as will be therapeutically effective and immunogenic. The quantity to be administered depends on the subject to be treated, capacity of the subject's immune system to synthesize antibodies, and the degree of protection desired. Precise amounts of active ingredient required to be administered depend on the judgment of the practitioner and are peculiar to each individual. However, suitable dosage ranges are of the order of several hundred micrograms active ingredient per individual. Suitable regimes for initial administration and booster shots are also variable, but are typified by an initial administration followed in one or two week intervals by a subsequent injection or other administration.

The recombinant expression constructs of the present invention are also useful in molecular biology to transform bacterial cells which do not ordinarily express a hemoglobin receptor protein to thereafter express this receptor. Such cells are

useful, *inter alia*, as intermediates for making cell membrane preparations useful for receptor binding activity assays, vaccine production, and the like, and in certain embodiments may themselves be used, *inter alia*, as vaccines or components of vaccines, as described above. The recombinant expression constructs of the present invention thus provide a method for screening potentially useful bactericidal and bacteriostatic drugs at advantageously lower cost than conventional screening protocols. While not completely eliminating the need for ultimate *in vivo* activity and toxicology assays, the constructs and cultures of the invention provide an important first screening step for the vast number of potentially useful bactericidal and bacteriostatic drugs synthesized, discovered or extracted from natural sources each year. In addition, such bactericidal or bacteriostatic drugs would be selected to utilize a nutritional pathway associated with infectious virulence in these types of bacteria, as disclosed in more detail below, thus selectively targeting bacteria associated with the development of serious infections *in vivo*.

Also, the invention provides both functional bacterial hemoglobin receptor proteins, membranes comprising such proteins, cells expressing such proteins, and the amino acid sequences of such proteins. This invention thereby provides sufficient structural and functional activity information to enable rational drug design of novel therapeutically-active antibacterial drugs using currently-available techniques (*see* Walters, "Computer-Assisted Modeling of Drugs", *in* Klegerman & Groves, eds., 1993, Pharmaceutical Biotechnology, Interpharm Press: Buffalo Grove, IL, pp. 165-174).

Nucleic acids and oligonucleotides of the present invention are useful as diagnostic tools for detecting the existence of a bacterial infection in a human, caused by a hemoglobin receptor protein-expressing pathological organism of *Neisseria* species. Such diagnostic reagents comprise nucleic acid hybridization probes of the invention and encompass paired oligonucleotide PCR primers, as described above. Methods provided by the invention include blot hybridization, *in situ* hybridization and *in vitro* amplification techniques for detecting the presence of pathogenic bacteria in a biological sample. Appropriate biological samples advantageously screened using the methods described herein include plasma, serum, lymph, cerebrospinal fluid, seminal fluid, mucosal tissue samples, biopsy samples, and other potential sites



of bacterial infection. It is also envisioned that the methods of the invention may be used to screen water, foodstuffs, pharmaceuticals, and other potential sources of infection.

5 The invention also provides antibodies that are immunologically reactive to a bacterial hemoglobin receptor protein or epitopes thereof provided by the invention. The antibodies provided by the invention may be raised, using methods well known in the art, in animals by inoculation with cells that express a bacterial hemoglobin receptor protein or epitopes thereof, cell membranes from such cells, whether crude membrane preparations or membranes purified using methods well known in the art, or purified preparations of proteins, including fusion proteins, particularly fusion proteins comprising epitopes of a bacterial hemoglobin receptor protein of the invention fused to heterologous proteins and expressed using genetic engineering means in bacterial, yeast or eukaryotic cells, said proteins being isolated from such cells to varying degrees of homogeneity using conventional biochemical means. Synthetic peptides made using established synthetic means *in vitro* and optionally conjugated with heterologous sequences of amino acids, are also encompassed in these methods to produce the antibodies of the invention. Animals that are used for such inoculations include individuals from species comprising cows, sheep, pigs, mice, rats, rabbits, hamsters, goats and primates. Preferred animals for inoculation are rodents (including mice, rats, hamsters) and rabbits. The most preferred animal is the mouse.

Cells that can be used for such inoculations, or for any of the other means used in the invention, include any cell that naturally expresses a bacterial hemoglobin receptor protein as provided by the invention, or any cell or cell line that expresses a bacterial hemoglobin receptor protein of the invention, or any epitope thereof, as a result of molecular or genetic engineering, or that has been treated to increase the expression of an endogenous or heterologous bacterial hemoglobin receptor protein by physical, biochemical or genetic means. Preferred cells are *E. coli* auxotrophic mutant *hema aroB* cells transformed with a recombinant expression construct of the invention and grown in media supplemented with hemin or hemoglobin as the sole iron (III) source, and cells of *Neisseria* species.

The present invention also provides monoclonal antibodies that are immunologically reactive with an epitope of a bacterial hemoglobin receptor protein of the invention, or fragment thereof, present on the surface of such cells, preferably *E. coli* cells. Such antibodies are made using methods and techniques well known to those of skill in the art. Monoclonal antibodies provided by the present invention are produced by hybridoma cell lines, that are also provided by the invention and that are made by methods well known in the art (*see* Harlow and Lane, 1988, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.).

Hybridoma cell lines are made by fusing individual cells of a myeloma cell line with spleen cells derived from animals immunized with a homogeneous preparation of a bacterial hemoglobin receptor protein, membranes comprised thereof, cells expressing such protein, or epitopes of a bacterial hemoglobin receptor protein, used *per se* or comprising a heterologous or fusion protein construct, as described above. The myeloma cell lines used in the invention include lines derived from myelomas of mice, rats, hamsters, primates and humans. Preferred myeloma cell lines are from mouse, and the most preferred mouse myeloma cell line is P3X63-Ag8.653. The animals from whom spleens are obtained after immunization are rats, mice and hamsters, preferably mice, most preferably Balb/c mice. Spleen cells and myeloma cells are fused using a number of methods well known in the art, including but not limited to incubation with inactivated Sendai virus and incubation in the presence of polyethylene glycol (PEG). The most preferred method for cell fusion is incubation in the presence of a solution of 45% (w/v) PEG-1450. Monoclonal antibodies produced by hybridoma cell lines can be harvested from cell culture supernatant fluids from *in vitro* cell growth; alternatively, hybridoma cells can be injected subcutaneously and/or into the peritoneal cavity of an animal, most preferably a mouse, and the monoclonal antibodies obtained from blood and/or ascites fluid.

Monoclonal antibodies provided by the present invention are also produced by recombinant genetic methods well known to those of skill in the art, and the present invention encompasses antibodies made by such methods that are immunologically reactive with an epitope of a bacterial hemoglobin receptor protein

of the invention. The present invention also encompasses fragments, including but not limited to F(ab) and F(ab)<sub>2</sub> fragments, of such antibody. Fragments are produced by any number of methods, including but not limited to proteolytic cleavage, chemical synthesis or preparation of such fragments by means of genetic engineering technology. The present invention also encompasses single-chain antibodies that are immunologically reactive with an epitope of a bacterial hemoglobin receptor protein, made by methods known to those of skill in the art.

The antibodies and fragments used herein can be labeled preferably with radioactive labels, by a variety of techniques. For example, the biologically active molecules can also be labeled with a radionucleotide via conjugation with the cyclic anhydride of diethylenetriamine penta-acetic acid (DPTA) or bromoacetyl aminobenzyl ethylamine diamine tetra-acidic acid (BABE). See Hnatowich *et al.* (1983, *Science* 220: 613-615) and Meares *et al.* (1984, *Anal. Biochem.* 142: 68-78, both references incorporated by reference) for further description of labeling techniques.

The present invention also encompasses an epitope of a bacterial hemoglobin receptor protein of the invention, comprised of sequences and/or a conformation of sequences present in the receptor molecule. This epitope may be naturally occurring, or may be the result of proteolytic cleavage of a receptor molecule and isolation of an epitope-containing peptide or may be obtained by synthesis of an epitope-containing peptide using methods well known to those skilled in the art. The present invention also encompasses epitope peptides produced as a result of genetic engineering technology and synthesized by genetically engineered prokaryotic or eukaryotic cells.

The invention also includes chimeric antibodies, comprised of light chain and heavy chain peptides immunologically reactive to a bacterial hemoglobin receptor protein-derived epitope. The chimeric antibodies embodied in the present invention include those that are derived from naturally occurring antibodies as well as chimeric antibodies made by means of genetic engineering technology well known to those of skill in the art.

Also provided by the present invention are diagnostic and therapeutic methods of detecting and treating an infection in a human, by a pathogenic organisms

expressing a bacterial hemoglobin receptor protein. Diagnostic reagents for use in such methods include the antibodies, most preferably monoclonal antibodies, of the invention. Such antibodies are used in conventional immunological techniques, including but not limited to enzyme-linked immunosorbent assay (ELISA),  
5 radioimmune assay (RIA), Western blot assay, immunological titration assays, immunological diffusion assays (such as the Ouchterlony assay), and others known to those of skill in the art. Also provided are epitopes derived from a bacterial hemoglobin receptor protein of the invention and immunologically cross-reactive to said antibodies, for use in any of the immunological techniques described herein.

10 Additional diagnostic assays include nucleic acid hybridization assays, using the nucleic acids of the invention or specifically-hybridizing fragments thereof, for sensitive detection of bacterial genomic DNA and/or mRNA. Such assays include various blot assays, such as Southern blots, Northern blots, dot blots, slot blots and the like, as well as *in vitro* amplification assays, such as the polymerase chain  
15 reaction assay (PCR), reverse transcriptase-polymerase chain reaction assay (RT-PCR), ligase chain reaction assay (LCR), and others known to those skilled in the art. Specific restriction endonuclease digestion of diagnostic fragments detected using any of the methods of the invention, analogous to restriction fragment linked polymorphism assays (RFLP) are also within the scope of this invention.

20 The invention also provides therapeutic methods and reagents for use in treating infections in a human, cause by a microorganism expressing a bacterial hemoglobin receptor protein of the invention, most preferably a bacteria of *Neisseria* species. Therapeutic reagents for use in such methods include the antibodies, most preferably monoclonal antibodies, of the invention, either *per se* or conjugated to  
25 bactericidal or bacteriostatic drugs or other antibiotic compounds effective against the infectious microorganism. In such embodiments, the antibodies of the invention comprise pharmaceutical compositions, additionally comprising appropriate pharmaceutically-acceptable carriers and adjuvants or other ancillary components where necessary. Suitable carriers are, for example, water, saline, dextrose,  
30 glycerol, ethanol, or the like and combinations thereof. In addition, if desired, the pharmaceutical formulation may contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, or other compounds which

enhance the effectiveness of the antibody. In these embodiments, it will be understood that the therapeutic agents of the invention serve to target the infectious bacteria, either by immunologically "tagging" the bacteria with an antibody of the invention for recognition by cytotoxic cells of a human's immune system, or by  
5 specifically delivering an antimicrobial drug to the infectious microorganism *via* the bacterial hemoglobin receptor protein.

Additional therapeutic reagents include the nucleic acids of the invention or fragments thereof, specifically antisense embodiments of such nucleic acids. Such antisense nucleic acids may be used themselves or embodied in a recombinant  
10 expression construct specific for antisense expression, wherein said construct is genetically engineered to co-opt a portion of the genome of a bacterial virus, preferably a bacteriophage, infectious for the bacterial pathogen responsible for the infection. In these embodiments, introduction of the antisense nucleic acids of the invention into the bacterial cell inhibits, attenuates or abolishes expression of the  
15 bacterial hemoglobin receptor, thereby reducing the virulence of the bacterial infection and enabling more effective antibacterial interventions. In additional embodiments, bacteriophage are provided bearing "knockout" copies of a bacterial hemoglobin receptor gene, whereby the phage achieves genetic mutation of the endogenous hemoglobin receptor gene in the infectious bacteria *via, for example,*  
20 homologous recombination of the exogenous knockout copy of the bacterial hemoglobin receptor gene with the endogenous hemoglobin receptor gene in the infectious microorganism.

The Examples which follow are illustrative of specific embodiments of the invention, and various uses thereof. They set forth for explanatory purposes only,  
25 and are not to be taken as limiting the invention.

### EXAMPLE 1

#### Plasmids, bacteria, and media

Plasmids and bacteria used herein are listed on Table 1. *E. coli* strains were  
30 routinely grown in Luria-Bertani (LB) broth supplemented with 5-aminolevulinic acid and 50mg/L hemin chloride as necessary. *N. meningitidis* 8013 is a serogroup C clinical isolate (Nassif *et al.*, 1993, *Mol. Microbiol.* 8: 719-725). The meningococci

were routinely grown on GCB agar (Difco) supplemented as described by Kellogg *et al.* (1963, *J. Bacteriol.* 85: 1274-1279), and incubated at 37°C under a 5% CO<sub>2</sub> atmosphere. Transformation of meningococci was performed as described by Nassif *et al.* (1992, *Mol. Microbiol.* 6: 591-597). When necessary, the following antibiotics  
5 were used with *E. coli*: rifampicin, 100 mg/L; tetracycline, 15 mg/L; kanamycin, 30 mg/L; chloramphenicol, 20 mg/L; carbenicillin, 100 mg/L. For *Neisseriae*, kanamycin at 100 mg/L was used when needed.

## EXAMPLE 2

### 10 Auxotroph Complementation Cloning of a hemoglobin Receptor Gene from *Neisseria meningitidis*

In order to identify *N. meningitidis* outer membrane receptor(s) involved in the uptake of haemin and/or haemoglobin iron, an auxotroph complementation cloning strategy was used, similar to the approach previously taken to identify the  
15 *Y. enterocolitica* and *V. cholerae* hemin receptors (*see* Stojiljkovic and Hantke, 1992, *EMBO J.* 11: 4359-4367; Henderson and Payne, 1994, *J. Bacteriol.* 176: 3269-3277). This strategy is based on the fact that the outer membrane of Gram-negative bacteria is impermeable to hemin (McConville and Charles, 1979, *J. Microbiol.* 113: 165-168) and therefore *E. coli* porphyrin biosynthesis mutants cannot grow on  
20 exogenously supplied hemin. If provided with the *N. meningitidis* outer membrane hemin receptor gene, the *E. coli* porphyrin mutant would be able to use exogenously supplied hemin as its porphyrin source.

A cosmid bank of *N. meningitidis* 8013 clone 6 DNA was prepared using conventional cosmid cloning methodologies (Sambrook *et al.*, 1989, *ibid.*). *N. meningitidis* bacterial DNA was partially digested by *MboI*, size fractionated on  
25 sucrose gradients and cloned into the *BamHI* site of the cosmid vector pLAFR2 (Riboli *et al.*, 1991, *Microb. Pathogen.* 10: 393-403). This cosmid bank was mobilized into the *E. coli* *hemA aroB Rif<sup>r</sup>* recipient strain by triparental matings using a conjugal plasmid pRK2013::Tn9. The mating mixture was plated on selective  
30 plates containing hemin chloride (50mg/L), 0.1 mM 2,2'-dipyridil and rifampicin (100 mg/L). Several clones growing on exogenously supplied haemin were isolated after an overnight incubation.

TABLE I

	<u>STRAIN</u>	<u>GENOTYPE</u>
5	<i>E. coli</i> K12	
	EB53	<i>hemA</i> , <i>aroB</i> , <i>rpoB</i>
	KP1041	MC4100 <i>tonB::Km'</i>
	H1388	<i>exbB::Tn10 Δlac pro</i>
	TSM348	<i>endA</i> , <i>hsdR</i> , <i>pro</i> , <i>supF</i> , pRK2013:: <i>Tn9</i>
10	IR754	EB53, <i>tonB::Km'</i>
	IR736	EB53, <i>exbB::Tn10</i>
	DH5α	<i>recA</i> , <i>gyrB</i>
	<i>N. meningitidis</i>	
	ATCC 13077	Serotype A
15	--	Serotype B*
	MC8013	clone 6, wild type
	MChmbR	<i>hmbR::aphA-3</i>
	<i>N. gonorrhoeae MS11A</i>	
20	<u>PLASMIDS</u>	
	pSUSK	pA15 replicon, chloramphenicol <sup>r</sup>
	pHEM22	pLAFR2, hemoglobin-utilizing cosmid
	pHEM44	pLAFR2, hemin-utilizing cosmid
	pIRS508	6kb <i>ClaI</i> , pSUSK
25	pIRS523	3kb <i>BamHI/SaII</i> , pUC19
	pIRS525	1.2kb <i>aphA-3</i> , in <i>NotI</i> site of pIRS523
	pIRS527	4kb <i>BamHI/ClaI</i> , pBluescript
	pIRS528	0.7kb <i>NotI/BamHI</i> , pBluescript
30	pIRS692	3.3kb <i>BamHI/HindIII</i> , SU(SK)

\* Laboratory collection

The hemin utilization phenotype of these transformants was tested by re-introduction of the cosmids into naive *E. coli hemA aroB* cells and by monitoring the growth on hemin-supplemented plates. The ability of *E. coli* strains to utilize heme or hemoglobin as the sole iron source was tested as previously described (Stojiljkovic and Hantke, 1992, *ibid.*). Cells were grown on LB agar supplemented with 50 $\mu$ M deferoxamine mesylate (an iron chelating agent, obtained from Sigma Chemical Co., St. Louis, MO). Filter discs (1/4 inches, Schleicher & Schuell, Inc., Keene, NH.) impregnated with the test compounds (20  $\mu$ L of 5 mg/ml stock solutions unless otherwise stated) were placed on these plates. After overnight growth at 37°C with 5% CO<sub>2</sub>, zones of growth around the discs were monitored. The iron-bound proteins tested in this assay (all obtained from Sigma Chemicals Co.) were hemoglobin from human, baboon, bovine and mouse sources, bovine hemin, human lactoferrin (90% iron saturated), and human transferrin (90% iron saturated, obtained from Boehringer Mannheim Biochemicals, Indianapolis, IN). A total of six hemin utilization positive cosmids were obtained using this protocol. Results using such assays are shown in Table II.

### EXAMPLE 3

#### Restriction Enzyme Digestion Mapping of Hemin Utilization Positive Cosmids

Cosmid DNA from six hemin-utilization positive cosmids obtained as described in Example 2 were digested with *Cla*I, and the resulting fragments were cloned into *Cla*I-digested pSU(SK) vector (obtained from Stratagene, LaJolla, CA). One subclone, containing a 6 kb *Cla*I fragment from cosmid cos22 (the resultant plasmid was designated pIRS508), was determined to allow utilization of hemin and hemoglobin by *E. coli hemA aroB* assayed as described in Example 2. Another such clone, containing an 11 kb *Cla*I fragment from cos44 was also determined to allow hemin utilization in these auxotrophic mutant cells. Restriction analysis and Southern hybridization indicated that the DNA fragments originating from cos22 and cos44 are unrelated.

The deduced restriction enzyme digestion map of cosmid clone pIRS508 is shown in Figure 1. Plasmid pIRS508 enabled *E. coli hemA aroB* to use both hemin and bovine hemoglobin as iron sources although growth on hemoglobin was



somewhat weaker than on hemin (Table II). Further subcloning localized the hemin/hemoglobin utilization locus to the *Bam*HI/*Hind*III fragment of the insert. In addition to sequences encoding the hemoglobin receptor gene (designated *hmbR*), sequences for a *Neisseria* insertion element (*IS1106*) and a portion of a *Neisseria* small repetitive element (*IR1*) are also represented in the Figure.

#### EXAMPLE 4

##### Nucleotide Sequence Analysis of a Cosmid Clone Encoding a *Neisseria* Hemoglobin Receptor Gene

The nucleotide sequence of the 3.3 kb *Bam*HI-*Hind*III DNA fragment carrying the *hmbR* gene and its promoter region was determined using the dideoxy chain termination method using a Sequenase 2.0 kit (obtained from U.S. Biochemicals, Cleveland, OH) and analyzed using a BioRad electrophoresis system, an AutoRead kit (obtained from Pharmacia, Uppsala, SE) and an ALF-370 automatic sequenator (Pharmacia, Uppsala, Sweden). Plasmid subclones for sequencing were produced by a nested deletion approach using Erase-a-Base kit (obtained from Promega Biotech, Madison, WI) using different restriction sites in the *hmbR* gene. The nucleotide and predicted amino acid sequences of the *hmbR* gene are shown in Figure 2

An open reading frame (ORF) encoding the *N. meningitidis*, serotype C hemoglobin receptor protein begins at position 470 of the sequence and encodes a protein having an amino acid sequence of 792 amino acids, with a calculated molecular weight of 85.5 kDa. A Shine-Delgarno sequence (SD) is found at position 460. The HmbR receptor protein contains a signal peptidase I recognition sequence at residues 22 to 24 of the protein (underlined), consistent with the fact that it is an outer membrane protein.

A typical Fur binding nucleotide sequence (designated "Fur box") was found in the promoter region of the *hmbR* gene (Figure 2). Like hemin utilization in *Yersiniae* and *Vibrio*, hemin and hemoglobin utilization in *Neisseria* are known to be iron-inducible phenotypes (West and Sparling, 1985, *Infect. Immun.* 47: 388-394; Dyer *et al.*, 1987, *Infect. Immun.* 55: 2171-2175). In Gram-negative bacteria, conditional expression of many iron utilization genes is regulated by the Fur

TABLE II

STRAIN	φ-TYPE	HEMIN IRON	PORPHYRIN	Hb IRON
<i>N. meningitidis</i>				
MC8013	wild type	+++	N.T.	+++
MChmbR	Hb <sup>R</sup> mutant	+++	N.T.	-
<i>E. coli</i>				
EB53	iron utilization	-	-	-
EB53 (pIRS508)	<i>tonB<sup>+</sup>, exbB<sup>+</sup>, hmbR<sup>+</sup></i>	+++	+++	+
IR754(pIRS508)	<i>tonB<sup>+</sup>, exbB<sup>+</sup>, hmbR<sup>+</sup></i>	-	-	-
IR736(pIRS508)	<i>tonB<sup>+</sup>, exbB<sup>+</sup>, hmbR<sup>+</sup></i>	-	-	-

N.T.-not tested. Use of hemin/hemoglobin as a porphyrin source was tested by scoring for growth of strains around hemin (5mg/mL) or hemoglobin (for *E. coli*, 10 mg/mL; for *N. meningitidis*, 5 mg/mL) discs on LB plates. The use of the hemin/hemoglobin as an iron source was tested similarly except NBD plates supplemented with 50 μL of 5 g/L delta-aminolevulinic acid were used (GCB plates supplemented with the 50μM Desferal in the case of *N. meningitidis*).

- : indicates no growth
- +: less than 100 mm of growth zone around the disc
- +++ : ±15 mm of growth zone around the disc.

repressor, which recognizes a 19 bp imperfect dyad repeat (Fur-box) in the promoter regions of Fur-repressed genes. Recently, a genetic screen (FURTA) for the identification of Fur-regulated genes from different Gram-negative bacteria was described (Stojiljkovic *et al.*, 1994, *J. Mol. Biol.* 236: 531-545), and this assay was used to test whether *hmbR* expression was controlled in this way. Briefly, a plasmid carrying a Fur-box sequence is transformed into an *E. coli* strain (H1717) which possesses a Fur-regulated *lac* fusion in the chromosome. Expression of this Fur-regulated *lac* fusion is normally repressed. Introduction of a multicopy Fur-box sequence on the plasmid titrates the available Fur repressor thus allowing expression of the Fur-regulated *lac* fusion (this phenotype is termed FURTA positive). Using this screen, the smallest insert fragment from cosmid pIRS508 that produced a FURTA positive result was a 0.7 kb *Bam*HI-*Not*I DNA fragment carried on plasmid pIRS528 (see Figure 1). This result indicated that the 0.7 kb *Bam*HI-*Not*I fragment carries a Fur-box and that gene expression from the *hmbR* promoter is controlled by a fur-type operon.

*N. meningitidis*, serotype C hemoglobin receptor protein was expressed *in vitro* using an *E. coli* S30 extract system from Promega Biotech (Madison, WI). The 3.3 kb *Bam*HI-*Hind*III fragment, expressed *in vitro*, encoded a 90kDa protein which corresponds in size to the predicted molecular weight of the unprocessed HmbR receptor. SDS/ 10% PAGE analysis showing the observed  $M_r$  of 90K is shown in Figure 3.

Immediately downstream of the *hmbR* gene (at positions 2955 to 3000 bp in Figure 2) was found a short nucleotide sequence that is 99% identical to the flanking sequence of the PIII gene of *N. gonorrhoeae* (Gotschlich *et al.*, 1987, *J. Exp. Med.* 165: 471-482). The first 26 bp of this sequence represents one half of the inverted repeat (IR1) of the *N. gonorrhoeae* small repetitive element. This element is found in approximately 20 copies in both *N. gonorrhoeae* and *N. meningitidis* (Correia *et al.*, 1988, *J. Biol. Chem.* 263: 12194-12198). The analysis of the nucleotide sequence from position 3027 to the *Cl*aI (3984) restriction site (only the nucleotide sequence from *Bam*HI (1) to *Hind*III (3370) is shown in Figure 2) indicated the presence of an IS1106 element (Knight *et al.*, 1992, *Mol. Microbiol.* 6: 1565-1573).

Interestingly, no nucleotide sequence similar to the IS1106 inverted repeat was found between the IR1 element and the beginning of the homology to IS1106.

These results were consistent with the cloning and identification of a novel hemoglobin receptor protein gene from *N. meningitidis*, embodied in a 3.3kb *Bam*HI/*Hind*III fragment of *N. meningitidis* genomic DNA.

## EXAMPLE 5

### Amino Acid Sequence Comparison of the *N. meningitidis* Hemoglobin Receptor Protein and *Neisseria* Lactoferrin and Transferrin Receptor Proteins

A comparison of the transferrin (Tbp1; Legrain *et al.*, 1993, *Gene* 130: 81-90), lactoferrin (LbpA; Pettersson *et al.*, 1993, *Infect. Immun.* 61: 4724-4733, and 1994, *J. Bacteriol.* 176: 1764-1766) and hemoglobin receptors (HmbR) from *N. meningitidis* is shown in Figure 4. The comparison was done with the CLASTAL program from the PC/GENE program package (Intelligenetics, Palo Alto, CA). Only the amino-terminal and carboxyl terminal segments of the proteins are shown. An asterisk indicates identity and a point indicates similarity at the amino acid level. Lactoferrin and transferrin receptors were found to share 44.4% identity in amino acid sequence. In contrast, homology between these proteins and the hemoglobin receptor disclosed herein was found to be significantly weaker (22% amino acid sequence identity with lactoferrin and 21% with transferrin receptor).

## EXAMPLE 6

### *TonB/ExbBD*-Dependence of Hemin Transport by the *N. meningitidis* Hemoglobin Receptor

It was known that the transport of iron-containing siderophores, some colicins and vitamin B12 across the outer membrane of *E. coli* depends on three cytoplasmic membrane proteins: TonB, ExbB and ExbD (Postle 1990, *Mol. Microbiol.* 133: 891-898; Braun and Hantke, 1991, in Winkelmann, (ed.), Handbook of Microbial Iron Chelates, CRC Press, Boca Raton, Fla., pp. 107-138). In *Yersinia* and *Hemophilus*, hemin uptake was shown to be a TonB-dependent process (Stojiljkovic and Hantke, 1992, *ibid.*; Jarosik *et al.*, 1994, *Infect. Immun.* 62: 2470-2477). Through direct interaction between the outer membrane receptors and the TonB cytoplasmic

machinery, the substrate bound to the receptor is internalized into the periplasm (Heller *et al.*, 1988, *Gene* 64: 147-153; Schoffler and Braun, 1989, *Molec. Gen. Genet.* 217: 378-383). This direct interaction has been associated with a particular amino acid sequence in membrane proteins associated with the TonB machinery.

5 All TonB-dependent receptors in Gram-negative bacteria contain several regions of high homology in their primary structures (Lundrigan and Kadner, 1986, *J. Biol. Chem.* 261: 10797-10801). In the amino acid sequence comparison described in Example 5, putative TonB-boxes of all three proteins are underlined. The carboxyl terminal end of the HmbR receptor contains the highly conserved  
10 terminal phenylalanine and position 782 arginine residues thought to be part of an outer membrane localization signal (Struyve *et al.*, 1991, *J. Mol. Biol.* 218: 141-148; Koebnik, 1993, *Trends Microbiol.* 1: 201). At residue 6 of the mature HmbR protein, an amino acid sequence - ETTPVKA - is similar in sequence to the so called TonB-boxes of several Gram-negative receptors (Heller *et al.*, 1988, *ibid.*).  
15 Interestingly, the putative TonB-box of HmbR has more homology to the TonB-box of the *N. gonorrhoeae* transferrin receptor (Cornelissen *et al.*, 1992, *J. Bacteriol.* 174: 5788-5797) than to the TonB-boxes of *E. coli* siderophore receptors. When the sequence of the HmbR receptor was compared with other TonB-dependent receptors, the highest similarity was found with *Y. enterocolitica* HemR receptor although the  
20 similarity was not as high as to the *Neisseria* receptors.

In order to prove the TonB-dependent nature of the *N. meningitidis*, serotype C hemoglobin receptor, *hmbR* was introduced into *exbB* and *tonB* mutants of *E. coli* EB53, and the ability of the strains to utilize hemin and hemoglobin as porphyrin and iron sources was assessed. In these assays, both mutants of *E. coli* EB53 were  
25 unable to use hemin either as a porphyrin source or as an iron source in the presence of a functional *hmbR* (Table 2). The usage of hemoglobin as an iron source was also affected (Table 2). These results are consistent with the notion that the *hmbR* gene product, the *N. meningitidis* hemoglobin receptor protein of the invention, is TonB-dependent, since expression of this gene in TonB wild type *E. coli* supported the use  
30 of hemin and hemoglobin as sole iron source in the experiments disclosed in Example 2.

## EXAMPLE 7

**Functional Demonstration that the *hmbR* Gene Product is the Hemoglobin Receptor Protein in *N. meningitidis***

As shown in the data presented in Table II, *hmbR* mediated both heme and hemoglobin utilization when expressed in *E. coli*, but hemoglobin utilization was less vigorous than heme utilization. To determine if the HmbR receptor has the same specificity in *N. meningitidis*, *hmbR* was inactivated with a 1.2kb kanamycin cassette (*aphA-3*; Nassif *et al.*, 1991, *ibid.*) and transformed into wild-type *N. meningitidis* 8013 clone 6 (serotype C) cells. The inactivation of the chromosomal *hmbR* copy of the Km-resistant transformants was confirmed by Southern hybridization, as shown in Figure 5. As can be seen from Figure 5, wild-type *N. meningitidis* genomic DNA contains only one copy of the *hmbR* gene (lanes 1 and 3). In the Km<sup>r</sup> transformants, the size of the DNA fragments containing the wild-type gene has increased by 1.2 kb, which is the size of the Kan cassette (Figure 5, lanes 2 and 4). When tested for its ability to utilize different iron-containing compounds, these mutant cells were found to be unable to use hemoglobin-bound iron, regardless of the source (human, bovine, baboon, mouse). The ability of the mutant to utilize hemoglobin-haptoglobin was not tested because the wild-type *N. meningitidis* strain is unable to use haptoglobin-haemoglobin complex as an iron source. However, the mutant was still able to use heme iron, lactoferrin- and transferrin-bound iron as well as citrate-iron (Table II). As the iron-containing component of hemoglobin is heme, a hemoglobin receptor would be expected to be capable of transporting heme into the periplasm. Indeed, the cloning strategy disclosed herein depended on the ability of the cloned meningococcal receptor to transport heme into the periplasm of *E. coli*. These results strongly suggest that *N. meningitidis* has at least two functional receptors that are involved in the internalization of heme-containing compounds. One is the hemoglobin receptor described herein, which allows the utilization of both heme and hemoglobin as iron sources. The other putative receptor in *N. meningitidis* is a heme receptor which allows utilization of only heme. This schema is also consistent with the isolation of several cosmid clones that allow *E. coli* EB53 to utilize heme. DNAs from these cosmids do not hybridize with our *hmbR* probe, indicating that these clones encode a structurally-distinct

receptor protein capable of transporting hemin into the periplasm of *N. meningitidis* cells.

### EXAMPLE 8

#### 5      Attenuation of Virulence in *hmbR* Mutant *N. meningitidis* Cells In Vivo

In order to test the importance of hemoglobin and hemin scavenging systems of *N. meningitidis* in vivo, the *hmbR* -mutant and the wild type strain of *N. meningitidis*, serotype C were inoculated into 5 day old infant rats and the numbers of bacteria recovered from blood and cerebrospinal fluid were followed. In these  
10 experiments, the method for the assessing *N. meningitidis*, serotype C virulence potential was essentially the same as described by Nassif *et al.* (1992, *ibid.*) using infant inbred Lewis rats (Charles River, Saint Aubin les Elbeufs, France). Inbred rats were used to minimize individual variations. Briefly, the 8013 strain was reactivated by 3 animal passages. After the third passage, bacteria were kept frozen  
15 in aliquots at -80° C. To avoid the possibility that modifications in the course of infection could result from selection of one spontaneous avirulent variant, one aliquot from the animal-passed frozen stock of 8013 was transformed with chromosomal DNA from the *hmbR* mutant, the resultant Kan<sup>r</sup> transformants were pooled without further purification and kept frozen at -80°C. For each experiment, all infant rats  
20 were from the same litter. *N. meningitidis* 8013 was grown overnight and 2 X 10<sup>6</sup> bacteria injected intraperitoneally into the infant rat. Three rats were used for each meningococcal strain. The course of infection was followed over a 24 hours time period with blood collected at the indicated times. At the 24 h time period, the rats were sacrificed, the cerebrospinal fluid (CSF) collected and the number of colony-forming units (CFU) determined. Each experiment was performed in replicate;  
25 similar results were obtained both times.

The results of these experiments are shown in Figure 6. The *hmbR* - strain, which is unable to use hemoglobin as an iron source, was recovered from the blood of infected animals in significantly lower numbers when compared with the wild type  
30 strain. Both the mutant and the wild type strain were still able to cross the blood-brain barrier as indicated by the isolation of bacteria from the cerebrospinal fluid.

These results indicate that hemoglobin represents an important iron source for *N. meningitidis* during growth *in vivo*.

#### EXAMPLE 9

##### 5 Polymerase Chain Reaction Amplification of Hemoglobin Receptor Genes from *N. meningitidis* Serotypes and *N. gonorrhoeae*

From the nucleotide sequence of the 3.3 kb *Bam*HI-*Hind*III DNA fragment carrying the *hmbR* gene and its promoter region was determined specific oligonucleotide primers for *in vitro* amplification of the homologous hemoglobin receptor protein genes from *N. meningitidis* serotypes A and B and *N. gonorrhoeae* MS11A as follows.

The following oligonucleotide primers were developed for *in vitro* amplification reactions using the polymerase chain reaction (PCR; Saiki *et al.*, 1988, *Science* 230: 1350-1354):

15 5'-AAACAGGTCTCGGCATAG-3' (sense primer) (SEQ ID No.:11)  
5'-CGCGAATTCAAACAGGTCTCGGCATAG-3' (SEQ ID No.:12)  
(antisense primer)

for amplifying the hemoglobin receptor protein from *N. meningitidis*, serotype A;

5'-CGCGAATTCAAAAACCTTCCATTCCAGCGATACG-3' (SEQ ID No.:13)  
20 (sense primer)

5'-TAAAACTTCCATTCCAGCGATACG-3' (antisense primer) (SEQ ID No.:14)

for amplifying the hemoglobin receptor protein from *N. meningitidis*, serotype B;

5'-AAACAGGTCTCGGCATAG-3' (sense primer) (SEQ ID No.:15)

or

25 5'-CGCGAATTCAAACAGGTCTCGGCATAG-3' (SEQ ID No.:16)  
(sense primer)

and

5'-CGCGAATTCAAAAACCTTCCATTCCAGCGATACG-3' (SEQ ID No.:17)  
(antisense primer)

30 or

5'-TAAAACTTCCATTCCAGCGATACG-3' (antisense primer) (SEQ ID No.:18)

for amplifying the hemoglobin receptor protein from *N. gonorrhoeae* MS11A.

Genomic DNA from *N. meningitidis* serotype A or B or *N. gonorrhoeae* species was prepared using standard techniques (see Sambrook, *et al.*, *ibid.*),  
35 including enzymatic degradation of bacterial cell walls, protoplast lysis, protease and RNase digestion, extraction with organic solvents such as phenol and/or chloroform,



and ethanol precipitation. Crude DNA preparations were also used. An amount (typically, about 0.1 $\mu$ g) of genomic DNA was used for each amplification reaction. A PCR amplification reaction consisted of *Pfu* polymerase (Stratagene, LaJolla, CA) and/or *Taq* polymerase (Boehringer Mannheim, Germany) in the appropriate buffer including about 20picomoles of each amplification primer and 200nanomoles of each deoxynucleoside triphosphate. Amplification reactions were performed according to the following scheme:

10	First cycle	5 min at 95°C
		2 min at 51°C
		6 min at 72°C
15	Cycles 2-13	45 sec at 95°C
		35 sec at 49°C
		10 min at 72°C
20	Cycles 14-30	25 sec at 95°C
		35 sec at 47°C
		10 min at 72°C

Upon completion of the amplification reaction, DNA fragments were cloned either blunt-ended or, after *Eco*RI digestion, into *Eco*RI digested pSUKS or pWKS30 vectors and transformed into bacteria. Positively-selected clones were then analyzed for the presence of recombinant inserts, which were sequenced as described above in Example 4.

As a result of these experiments, three clones encoding the hemoglobin receptor genes from *N. meningitidis* serotypes A and B and *N. gonorrhoeae* MS11A were cloned and the sequence of these genes determined. The nucleic acid sequence for each of these genes are shown in Figures 7 (*N. meningitidis*, serotype A), 8 (*N. meningitidis*, serotype A) and 9 (*N. gonorrhoeae* MS11A).

The degree of homology between the cloned hemoglobin receptors from the different *N. meningitidis* serotypes and *N. gonorrhoeae* MS11A was assessed by nucleic acid and amino acid sequence comparison, as described in Example 5 above. The results of these comparisons are shown in Figures 10 and 11, respectively.

Hemoglobin receptor genes from the three *N. meningitidis* serotypes and *N. gonorrhoeae* MS11A were found to be from 86.5% to 93.4% homologous; the most homologous nucleic acids were *N. meningitidis* serotypes B and C, and the most divergent nucleic acids were *N. meningitidis* serotype B and *N. gonorrhoeae* MS11A (Figure 10 and Table III). Hemoglobin receptor proteins from all four *Neisseria* species showed a high degree of homology to the other members of the group, ranging from 87% homology between the hemoglobin receptor proteins from *N. gonorrhoeae* MS11A and *N. meningitidis* serotype B to 93% homology between hemoglobin receptor proteins from *N. meningitidis* serotypes A and B (Figure 11). In this comparison, all four receptors were found to share 84.7% amino acid sequence identity, and up to 11.6% sequence similarity (*i.e.*, chemically-related amino acid residues at homologous sites within the amino acid sequence). The non-conserved amino acids were found clustered in the regions of the amino acid sequence corresponding to the external loops in the predicted topographical structure of the hemoglobin receptor proteins.

It should be understood that the foregoing disclosure emphasizes certain specific embodiments of the invention and that all modifications or alternatives equivalent thereto are within the spirit and scope of the invention as set forth in the appended claims.

TABLE III

*	A			B			C			MS11		
A	X			92.2%			93.0%			90.4%		
B	93.3%			X			93.4%			86.5%		
C	93.2%			93%			X			90.4%		
MS11	91.1%			86.8%			91.4%			X		

\* The numbers in the upper quadrant of the Table (in boldface) represent nucleic acid sequence homology between the different hemoglobin receptor genes of the invention, while the numbers in the lower quadrant of the Table represent amino acid sequence homology between the different hemoglobin receptor proteins

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

(A) NAME: Oregon Health Sciences University  
 (B) STREET: 3181 S.W. Sam Jackson Park Road  
 (C) CITY: Portland  
 (D) STATE: Oregon  
 (E) COUNTRY: USA  
 (F) POSTAL CODE (ZIP): 97201-3098  
 (G) TELEPHONE: 503-494-8200  
 (H) TELEFAX: (503)-494-4729

(ii) TITLE OF INVENTION: A Novel Bacterial Hemoglobin Receptor  
 and Uses

(iii) NUMBER OF SEQUENCES: 18

## (iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
 (B) COMPUTER: IBM PC compatible  
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

## (v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2373 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

(A) NAME/KEY: CDS  
 (B) LOCATION: 1..2373

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG	AAA	CCA	TTA	CAA	ATG	CTC	CCT	ATC	GCC	GCG	CTG	GTC	GGC	AGT	ATT
Met	Lys	Pro	Leu	Gln	Met	Leu	Pro	Ile	Ala	Ala	Leu	Val	Gly	Ser	Ile
1				5					10					15	

48

TTC GGC AAT CCG GTC TTG GCA GCA GAT GAA GCT GCA ACT GAA ACC ACA Phe Gly Asn Pro Val Leu Ala Ala Asp Glu Ala Ala Thr Glu Thr Thr	96
20 25 30	
CCC GTT AAG GCA GAG ATA AAA GCA GTG CGC GTT AAA GGT CAG CGC AAT Pro Val Lys Ala Glu Ile Lys Ala Val Arg Val Lys Gly Gln Arg Asn	144
35 40 45	
GCG CCT GCG GCT GTG GAA CGC GTC AAC CTT AAC CGT ATC AAA CAA GAA Ala Pro Ala Ala Val Glu Arg Val Asn Leu Asn Arg Ile Lys Gln Glu	192
50 55 60	
ATG ATA CGC GAC AAT AAA GAC TTG GTG CGC TAT TCC ACC GAT GTC GGC Met Ile Arg Asp Asn Lys Asp Leu Val Arg Tyr Ser Thr Asp Val Gly	240
65 70 75 80	
TTG AGC GAC AGC GGC CGC CAT CAA AAA GGC TTT GCT GTT CGC GGC GTG Leu Ser Asp Ser Gly Arg His Gln Lys Gly Phe Ala Val Arg Gly Val	288
85 90 95	
GAA GGC AAC CGT GTC GGC GTG AGC ATA GAC GGT GTA AAC CTG CCT GAT Glu Gly Asn Arg Val Gly Val Ser Ile Asp Gly Val Asn Leu Pro Asp	336
100 105 110	
TCC GAA GAA AAC TCG CTG TAC GCC CGT TAT GGC AAC TTC AAC AGC TCG Ser Glu Glu Asn Ser Leu Tyr Ala Arg Tyr Gly Asn Phe Asn Ser Ser	384
115 120 125	
CGT TTG TCT ATC GAC CCC GAA CTC GTA CGC AAT ATT GAA ATC GTG AAG Arg Leu Ser Ile Asp Pro Glu Leu Val Arg Asn Ile Glu Ile Val Lys	432
130 135 140	
GGC GCA GAC TCT TTC AAT ACC GGC AGT GGT GCA TTG GGC GGC GGT GTG Gly Ala Asp Ser Phe Asn Thr Gly Ser Gly Ala Leu Gly Gly Gly Val	480
145 150 155 160	
AAT TAC CAA ACG CTG CAA GGC CGT GAT TTG CTG TTG GAC GAC AGG CAA Asn Tyr Gln Thr Leu Gln Gly Arg Asp Leu Leu Leu Asp Asp Arg Gln	528
165 170 175	
TTC GGC GTG ATG ATG AAA AAC GGT TAC AGC ACG CGT AAC CGT GAA TGG Phe Gly Val Met Met Lys Asn Gly Tyr Ser Thr Arg Asn Arg Glu Trp	576
180 185 190	
ACA AAT ACC CTC GGT TTC GGT GTG AGT AAC GAC CGC GTG GAT GCT GCT Thr Asn Thr Leu Gly Phe Gly Val Ser Asn Asp Arg Val Asp Ala Ala	624
195 200 205	
TTG CTG TAT TCG CAA CGG CGC GGC CAT GAA ACC GAA AGC GCG GGC AAC Leu Leu Tyr Ser Gln Arg Arg Gly His Glu Thr Glu Ser Ala Gly Asn	672
210 215 220	
CGC GGC TAT CCG GTA GAA GGT GCG GGT AAA GAA ACG AAT ATC CGC GGT Arg Gly Tyr Pro Val Glu Gly Ala Gly Lys Glu Thr Asn Ile Arg Gly	720
225 230 235 240	
TCC GCC CGC GGC ATC CCC GAT CCG TCC AAA CAC AAA TAC CAC AAC TTC Ser Ala Arg Gly Ile Pro Asp Pro Ser Lys His Lys Tyr His Asn Phe	768
245 250 255	
TTG GGT AAG ATT GCT TAT CAA ATC AAC GAC AAC CAC CGC ATC GGC GCA Leu Gly Lys Ile Ala Tyr Gln Ile Asn Asp Asn His Arg Ile Gly Ala	816
260 265 270	
TCG CTC AAC GGT CAG CAG GGG CAT AAT TAC ACG GTT GAA GAG TCT TAT Ser Leu Asn Gly Gln Gln Gly His Asn Tyr Thr Val Glu Glu Ser Tyr	864
275 280 285	

AAC Asn 290	CTG Leu	ACC Thr	GCT Ala	TCT Ser	TCC Ser	TGG Trp 295	CGC Arg	GAA Glu	GCC Ala	GAT Asp	GAC Asp 300	GTA Val	AAC Asn	AGA Arg	CGG Arg	912
CGC Arg 305	AAT Asn	GCC Ala	AAC Asn	CTC Leu	TTT Phe 310	TAC Tyr	GAA Glu	TGG Trp	ATG Met	CCT Pro 315	GAT Asp	TCA Ser	AAT Asn	TGG Trp	TTG Leu 320	960
TCG Ser	TCT Ser	TTG Leu	AAG Lys	GCG Ala 325	GAC Asp	TTC Phe	GAT Asp	TAT Tyr	CAG Gln 330	AAA Lys	ACC Thr	AAA Lys	GTG Val	GCG Ala 335	GCG Ala	1008
ATT Ile	AAC Asn	AAA Lys	GGT Gly 340	TCG Ser	TTC Phe	CCG Pro	ACG Thr	AAT Asn 345	TAC Tyr	ACC Thr	ACA Thr	TGG Trp	GAA Glu 350	ACT Thr	GAG Glu	1056
TAC Tyr	CAT His	AAA Lys 355	AAG Lys	GAA Glu	GTT Val	GGC Gly	GAA Glu 360	ATA Ile	TAC Tyr	AAC Asn	CGC Arg	AGC Ser 365	ATG Met	GAC Asp	ACC Thr	1104
CGA Arg 370	TTC Phe	AAA Lys	CGT Arg	TTT Phe	ACT Thr	TTG Leu 375	CGT Arg	TTG Leu	GAC Asp	AGC Ser	CAT His 380	CCG Pro	TTG Leu	CAA Gln	CTC Leu	1152
GGG Gly 385	GGG Gly	GGG Gly	CGA Arg	CAC His	CGC Arg 390	CTG Leu	TCG Ser	TTT Phe	AAA Lys	ACT Thr 395	TTC Phe	GCC Ala	AGC Ser	CGC Arg	CGT Arg 400	1200
GAT Asp	TTT Phe	GAA Glu	AAC Asn 405	CTA Leu	AAC Asn	CGC Arg	GAC Asp	GAT Asp	TAT Tyr 410	TAC Tyr	TTC Phe	AGC Ser	GGC Gly	CGT Arg 415	GTT Val	1248
GTT Val	CGA Arg	ACC Thr	ACC Thr 420	AGC Ser	AGT Ser	ATC Ile	CAG Gln	CAT His 425	CCG Pro	GTG Val	AAA Lys	ACC Thr	ACC Thr	AAC Asn	TAC Tyr	1296
GGT Gly	TTC Phe	TCA Ser 435	CTG Leu	TCT Ser	GAC Asp	CAA Gln	ATT Ile 440	CAA Gln	TGG Trp	AAC Asn	GAC Asp	GTG Val 445	TTC Phe	AGT Ser	AGC Ser	1344
CGC Arg 450	GCA Ala	GGT Gly	ATC Ile	CGT Arg	TAC Tyr	GAC Asp 455	CAC His	ACC Thr	AAA Lys	ATG Met	ACG Thr 460	CCT Pro	CAG Gln	GAA Glu	TTG Leu	1392
AAT Asn 465	GCC Ala	GAG Glu	TGT Cys	CAT His	GCT Ala 470	TGT Cys	GAC Asp	AAA Lys	ACA Thr	CCA Pro 475	CCT Pro	GCA Ala	GCC Ala	AAC Asn	ACT Thr 480	1440
TAT Tyr	AAA Lys	GGC Gly	TGG Trp	AGC Ser	GGT Gly 485	TTT Phe	GTC Val	GGC Gly	TTG Leu 490	GCG Ala	GCG Ala	CAA Gln	CTG Leu	AAT Asn	CAG Gln 495	1488
GCT Ala	TGG Trp	CGT Arg	GTC Val 500	GGT Gly	TAC Tyr	GAC Asp	ATT Ile	ACT Thr 505	TCC Ser	GGC Gly	TAC Tyr	CGT Arg	GTC Val 510	CCC Pro	AAT Asn	1536
GCG Ala	TCC Ser	GAA Glu 515	GTG Val	TAT Tyr	TTC Phe	ACT Thr 520	TAC Tyr	AAC Asn	CAC His	GGT Gly	TCG Ser 525	GGT Asn	AAT Asn	TGG Trp	CTG Leu	1584
CCC Pro 530	AAT Pro	CCC Pro	AAC Asn	CTG Leu	AAA Lys 535	GCC Ala	GAG Glu	CGC Arg	AGC Ser	ACC Thr	ACC Thr 540	CAC His	ACC Thr	CTG Leu	TCT Ser	1632
CTG Leu 545	CAA Gln	GGC Gly	CGC Arg	AGC Ser	GAA Glu 550	AAA Lys	GGC Gly	ATG Met	CTG Leu	GAT Asp 555	GCC Ala	AAC Asn	CTG Leu	TAT Tyr	CAA Gln 560	1680

AGC AAT TAC CGC AAT TTC CTG TCT GAA GAG CAG AAG CTG ACC ACC AGC Ser Asn Tyr Arg Asn Phe Leu Ser Glu Glu Gln Lys Leu Thr Thr Ser 565 570 575	1728
GGC ACT CCC GGC TGT ACT GAG GAA AAT GCT TAC TAC AGT ATA TGC AGC Gly Thr Pro Gly Cys Thr Glu Glu Asn Ala Tyr Tyr Ser Ile Cys Ser 580 585 590	1776
GAC CCC TAC AAA GAA AAA CTG GAT TGG CAG ATG AAA AAT ATC GAC AAG Asp Pro Tyr Lys Glu Lys Leu Asp Trp Gln Met Lys Asn Ile Asp Lys 595 600 605	1824
GCC AGA ATC CGC GGT ATC GAG CTG ACA GGC CGT CTG AAT GTG GAC AAA Ala Arg Ile Arg Gly Ile Glu Leu Thr Gly Arg Leu Asn Val Asp Lys 610 615 620	1872
GTA GCG TCT TTT GTT CCT GAG GGC TGG AAA CTG TTC GGC TCG CTG GGT Val Ala Ser Phe Val Pro Glu Gly Trp Lys Leu Phe Gly Ser Leu Gly 625 630 635 640	1920
TAT GCG AAA AGC AAA CTG TCG GGC GAC AAC AGC CTG CTG TCC ACA CAG Tyr Ala Lys Ser Lys Leu Ser Gly Asp Asn Ser Leu Leu Ser Thr Gln 645 650 655	1968
CCG CTG AAA GTG ATT GCC GGT ATC GAC TAT GAA AGT CCG AGC GAA AAA Pro Leu Lys Val Ile Ala Gly Ile Asp Tyr Glu Ser Pro Ser Glu Lys 660 665 670	2016
TGG GGC GTA TTC TCC CGC CTG ACC TAT CTG GGC GCG AAA AAG GTC AAA Trp Gly Val Phe Ser Arg Leu Thr Tyr Leu Gly Ala Lys Lys Val Lys 675 680 685	2064
GAC GCG CAA TAC ACC GTT TAT GAA AAC AAG GGC TGG GGT ACG CCT TTG Asp Ala Gln Tyr Thr Val Tyr Glu Asn Lys Gly Trp Gly Thr Pro Leu 690 695 700	2112
CAG AAA AAG GTA AAA GAT TAC CCG TGG CTG AAC AAG TCG GCT TAT GTG Gln Lys Lys Val Lys Asp Tyr Pro Trp Leu Asn Lys Ser Ala Tyr Val 705 710 715 720	2160
TTC GAT ATG TAC GGC TTC TAC AAA CCG GTG AAA AAC CTG ACC CTG CGT Phe Asp Met Tyr Gly Phe Tyr Lys Pro Val Lys Asn Leu Thr Leu Arg 725 730 735	2208
GCG GGC GTG TAC AAC CTG TTC AAC CGC AAA TAC ACC ACT TGG GAT TCC Ala Gly Val Tyr Asn Leu Phe Asn Arg Lys Tyr Thr Thr Trp Asp Ser 740 745 750	2256
CTG CGC GGT TTA TAT AGC TAC AGC ACC ACC AAT GCG GTC GAC CGC GAT Leu Arg Gly Leu Tyr Ser Tyr Ser Thr Thr Asn Ala Val Asp Arg Asp 755 760 765	2304
GGC AAA GGC TTA GAT CGC TAC CGC GCC CCA GGC CGC AAT TAC GCC GTA Gly Lys Gly Leu Asp Arg Tyr Arg Ala Pro Gly Arg Asn Tyr Ala Val 770 775 780	2352
TCG CTG GAA TGG AAG TTT TAA Ser Leu Glu Trp Lys Phe * 785 790	2373

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 790 amino acids  
(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Lys Pro Leu Gln Met Leu Pro Ile Ala Ala Leu Val Gly Ser Ile
 1           5           10           15
Phe Gly Asn Pro Val Leu Ala Ala Asp Glu Ala Ala Thr Glu Thr Thr
          20           25           30
Pro Val Lys Ala Glu Ile Lys Ala Val Arg Val Lys Gly Gln Arg Asn
          35           40           45
Ala Pro Ala Ala Val Glu Arg Val Asn Leu Asn Arg Ile Lys Gln Glu
          50           55           60
Met Ile Arg Asp Asn Lys Asp Leu Val Arg Tyr Ser Thr Asp Val Gly
 65           70           75           80
Leu Ser Asp Ser Gly Arg His Gln Lys Gly Phe Ala Val Arg Gly Val
          85           90           95
Glu Gly Asn Arg Val Gly Val Ser Ile Asp Gly Val Asn Leu Pro Asp
          100          105          110
Ser Glu Glu Asn Ser Leu Tyr Ala Arg Tyr Gly Asn Phe Asn Ser Ser
          115          120          125
Arg Leu Ser Ile Asp Pro Glu Leu Val Arg Asn Ile Glu Ile Val Lys
          130          135          140
Gly Ala Asp Ser Phe Asn Thr Gly Ser Gly Ala Leu Gly Gly Gly Val
          145          150          155          160
Asn Tyr Gln Thr Leu Gln Gly Arg Asp Leu Leu Leu Asp Asp Arg Gln
          165          170          175
Phe Gly Val Met Met Lys Asn Gly Tyr Ser Thr Arg Asn Arg Glu Trp
          180          185          190
Thr Asn Thr Leu Gly Phe Gly Val Ser Asn Asp Arg Val Asp Ala Ala
          195          200          205
Leu Leu Tyr Ser Gln Arg Arg Gly His Glu Thr Glu Ser Ala Gly Asn
          210          215          220
Arg Gly Tyr Pro Val Glu Gly Ala Gly Lys Glu Thr Asn Ile Arg Gly
          225          230          235          240
Ser Ala Arg Gly Ile Pro Asp Pro Ser Lys His Lys Tyr His Asn Phe
          245          250          255
Leu Gly Lys Ile Ala Tyr Gln Ile Asn Asp Asn His Arg Ile Gly Ala
          260          265          270
Ser Leu Asn Gly Gln Gln Gly His Asn Tyr Thr Val Glu Glu Ser Tyr
          275          280          285
Asn Leu Thr Ala Ser Ser Trp Arg Glu Ala Asp Asp Val Asn Arg Arg
          290          295          300
Arg Asn Ala Asn Leu Phe Tyr Glu Trp Met Pro Asp Ser Asn Trp Leu
          305          310          315          320

```



Ser Ser Leu Lys Ala Asp Phe Asp Tyr Gln Lys Thr Lys Val Ala Ala  
 325 330 335  
 Ile Asn Lys Gly Ser Phe Pro Thr Asn Tyr Thr Thr Trp Glu Thr Glu  
 340 345 350  
 Tyr His Lys Lys Glu Val Gly Glu Ile Tyr Asn Arg Ser Met Asp Thr  
 355 360 365  
 Arg Phe Lys Arg Phe Thr Leu Arg Leu Asp Ser His Pro Leu Gln Leu  
 370 375 380  
 Gly Gly Gly Arg His Arg Leu Ser Phe Lys Thr Phe Ala Ser Arg Arg  
 385 390 395 400  
 Asp Phe Glu Asn Leu Asn Arg Asp Asp Tyr Tyr Phe Ser Gly Arg Val  
 405 410 415  
 Val Arg Thr Thr Ser Ser Ile Gln His Pro Val Lys Thr Thr Asn Tyr  
 420 425 430  
 Gly Phe Ser Leu Ser Asp Gln Ile Gln Trp Asn Asp Val Phe Ser Ser  
 435 440 445  
 Arg Ala Gly Ile Arg Tyr Asp His Thr Lys Met Thr Pro Gln Glu Leu  
 450 455 460  
 Asn Ala Glu Cys His Ala Cys Asp Lys Thr Pro Pro Ala Ala Asn Thr  
 465 470 475 480  
 Tyr Lys Gly Trp Ser Gly Phe Val Gly Leu Ala Ala Gln Leu Asn Gln  
 485 490 495  
 Ala Trp Arg Val Gly Tyr Asp Ile Thr Ser Gly Tyr Arg Val Pro Asn  
 500 505 510  
 Ala Ser Glu Val Tyr Phe Thr Tyr Asn His Gly Ser Gly Asn Trp Leu  
 515 520 525  
 Pro Asn Pro Asn Leu Lys Ala Glu Arg Ser Thr Thr His Thr Leu Ser  
 530 535 540  
 Leu Gln Gly Arg Ser Glu Lys Gly Met Leu Asp Ala Asn Leu Tyr Gln  
 545 550 555 560  
 Ser Asn Tyr Arg Asn Phe Leu Ser Glu Glu Gln Lys Leu Thr Thr Ser  
 565 570 575  
 Gly Thr Pro Gly Cys Thr Glu Glu Asn Ala Tyr Tyr Ser Ile Cys Ser  
 580 585 590  
 Asp Pro Tyr Lys Glu Lys Leu Asp Trp Gln Met Lys Asn Ile Asp Lys  
 595 600 605  
 Ala Arg Ile Arg Gly Ile Glu Leu Thr Gly Arg Leu Asn Val Asp Lys  
 610 615 620  
 Val Ala Ser Phe Val Pro Glu Gly Trp Lys Leu Phe Gly Ser Leu Gly  
 625 630 635 640  
 Tyr Ala Lys Ser Lys Leu Ser Gly Asp Asn Ser Leu Leu Ser Thr Gln  
 645 650 655  
 Pro Leu Lys Val Ile Ala Gly Ile Asp Tyr Glu Ser Pro Ser Glu Lys  
 660 665 670

Trp Gly Val Phe Ser Arg Leu Thr Tyr Leu Gly Ala Lys Lys Val Lys  
 675 680 685  
 Asp Ala Gln Tyr Thr Val Tyr Glu Asn Lys Gly Trp Gly Thr Pro Leu  
 690 695 700  
 Gln Lys Lys Val Lys Asp Tyr Pro Trp Leu Asn Lys Ser Ala Tyr Val  
 705 710 715 720  
 Phe Asp Met Tyr Gly Phe Tyr Lys Pro Val Lys Asn Leu Thr Leu Arg  
 725 730 735  
 Ala Gly Val Tyr Asn Leu Phe Asn Arg Lys Tyr Thr Thr Trp Asp Ser  
 740 745 750  
 Leu Arg Gly Leu Tyr Ser Tyr Ser Thr Thr Asn Ala Val Asp Arg Asp  
 755 760 765  
 Gly Lys Gly Leu Asp Arg Tyr Arg Ala Pro Gly Arg Asn Tyr Ala Val  
 770 775 780  
 Ser Leu Glu Trp Lys Phe  
 785 790

## (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2375 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..2375

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG AAA CCA TTA CAA ATG CCC CCT ATC GCC GCG CTG CTC GGC AGT ATT	48
Met Lys Pro Leu Gln Met Pro Pro Ile Ala Ala Leu Leu Gly Ser Ile	
1 5 10 15	
TTC GGC AAT CCG GTC TTT GCG GCA GAT GAA GCT GCA ACT GAA ACC ACA	96
Phe Gly Asn Pro Val Phe Ala Ala Asp Glu Ala Ala Thr Glu Thr Thr	
20 25 30	
CCC GTT AAG GCA GAG GTA AAA GCA GTG CGC GTT AAA GGT CAG CGC AAT	144
Pro Val Lys Ala Glu Val Lys Ala Val Arg Val Lys Gly Gln Arg Asn	
35 40 45	
GCG CCT GCG GCT GTG GAA CGC GTC AAC CTT AAC CGT ATC AAA CAA GAA	192
Ala Pro Ala Ala Val Glu Arg Val Asn Leu Asn Arg Ile Lys Gln Glu	
50 55 60	
ATG ATA CGC GAC AAT AAA GAC TTG GTG CGC TAT TCC ACC GAT GTC GGC	240
Met Ile Arg Asp Asn Lys Asp Leu Val Arg Tyr Ser Thr Asp Val Gly	
65 70 75 80	
TTG AGC GAC AGG AGC CGT CAT CAA AAA GGC TTT GCC ATT CGC GGC GTG	288
Leu Ser Asp Arg Ser Arg His Gln Lys Gly Phe Ala Ile Arg Gly Val	
85 90 95	

GAA GGC GAC CGT GTC GGC GTT AGT ATT GAC GGC GTA AAC CTG CCT GAT Glu Gly Asp Arg Val Gly Val Ser Ile Asp Gly Val Asn Leu Pro Asp 100 105 110	336
TCC GAA GAA AAC TCG CTG TAC GCC CGT TAT GGC AAC TTC AAC AGC TCG Ser Glu Glu Asn Ser Leu Tyr Ala Arg Tyr Gly Asn Phe Asn Ser Ser 115 120 125	384
CGT CTG TCT ATC GAC CCC GAA CTC GTG CGC AAC ATC GAC ATC GTA AAA Arg Leu Ser Ile Asp Pro Glu Leu Val Arg Asn Ile Asp Ile Val Lys 130 135 140	432
GGG GCG GAC TCT TTC AAT ACC GGC AGC GGC GCC TTG GGC GGC GGT GTG Gly Ala Asp Ser Phe Asn Thr Gly Ser Gly Ala Leu Gly Gly Gly Val 145 150 155 160	480
AAT TAC CAA ACC CTG CAA GGA CGT GAC TTA CTG TTG CCT GAA CGG CAG Asn Tyr Gln Thr Leu Gln Gly Arg Asp Leu Leu Leu Pro Glu Arg Gln 165 170 175	528
TTC GGC GTG ATG ATG AAA AAC GGT TAC AGC ACG CGT AAC CGT GAA TGG Phe Gly Val Met Met Lys Asn Gly Tyr Ser Thr Arg Asn Arg Glu Trp 180 185 190	576
ACA AAT ACC CTC GGT TTC GGC GTG AGC AAC GAC CGC GTG GAT GCC GCT Thr Asn Thr Leu Gly Phe Gly Val Ser Asn Asp Arg Val Asp Ala Ala 195 200 205	624
TTG CTG TAT TCG CAA CGG CGC GGC CAT GAA ACT GAA AGC GCG GGC AAG Leu Leu Tyr Ser Gln Arg Arg Gly His Glu Thr Glu Ser Ala Gly Lys 210 215 220	672
CGT GGT TAT CCG GTA GAG GGT GCT GGT AGC GGA GCG AAT ATC CGT GGT Arg Gly Tyr Pro Val Glu Gly Ala Gly Ser Gly Ala Asn Ile Arg Gly 225 230 235 240	720
TCT GCG CGC GGT ATT CCT GAT CCG TCC CAA CAC AAA TAC CAC AGC TTC Ser Ala Arg Gly Ile Pro Asp Pro Ser Gln His Lys Tyr His Ser Phe 245 250 255	768
TTG GGT AAG ATT GCT TAT CAA ATC AAC GAC AAC CAC CGC ATC GGC GCA Leu Gly Lys Ile Ala Tyr Gln Ile Asn Asp Asn His Arg Ile Gly Ala 260 265 270	816
TCG CTC AAC GGT CAG CAG GGG CAT AAT TAC ACG GTT GAA GAG TCT TAC Ser Leu Asn Gly Gln Gln Gly His Asn Tyr Thr Val Glu Glu Ser Tyr 275 280 285	864
AAC CTG CTT GCT TCT TAT TGG CGT GAA GCT GAC GAT GTC AAC AGA CGG Asn Leu Leu Ala Ser Tyr Trp Arg Glu Ala Asp Asp Val Asn Arg Arg 290 295 300	912
CGT AAC ACC AAC CTC TTT TAC GAA TGG ACG CCG GAA TCC GAC CGG TTG Arg Asn Thr Asn Leu Phe Tyr Glu Trp Thr Pro Glu Ser Asp Arg Leu 305 310 315 320	960
TCT ATG GTA AAA GCG GAT GTC GAT TAT CAA AAA ACC AAA GTA TCT GCG Ser Met Val Lys Ala Asp Val Asp Tyr Gln Lys Thr Lys Val Ser Ala 325 330 335	1008
GTC AAC TAC AAA GGT TCG TTC CCG ACG AAT TAC ACC ACA TGG GAA ACC Val Asn Tyr Lys Gly Ser Phe Pro Thr Asn Tyr Thr Thr Trp Glu Thr 340 345 350	1056
GAG TAC CAT AAA AAG GAA GTT GGC GAA ATC TAT AAC CGC AGC ATG GAT Glu Tyr His Lys Lys Glu Val Gly Glu Ile Tyr Asn Arg Ser Met Asp 355 360 365	1104

ACA ACC TTC AAA CGT ATT ACG CTG CGT ATG GAC AGC CAT CCG TTG CAA Thr Thr Phe Lys Arg Ile Thr Leu Arg Met Asp Ser His Pro Leu Gln 370 375 380	1152
CTC GGG GGG GGG CGA CAC CGC CTG TCG TTC AAA ACC TTT GCC GGG CAG Leu Gly Gly Gly Arg His Arg Leu Ser Phe Lys Thr Phe Ala Gly Gln 385 390 395 400	1200
CGT GAT TTT GAA AAC TTA AAC CGC GAC GAT TAC TAC TTC AGC GGC CGT Arg Asp Phe Glu Asn Leu Asn Arg Asp Asp Tyr Tyr Phe Ser Gly Arg 405 410 415	1248
GTT GTT CGA ACC ACC AAC AGT ATC CAG CAT CCG GTG AAA ACC ACC AAC Val Val Arg Thr Thr Asn Ser Ile Gln His Pro Val Lys Thr Thr Asn 420 425 430	1296
TAC GGT TTC TCG CTG TCC GAC CAA ATC CAA TGG AAC GAC GTG TTC AGT Tyr Gly Phe Ser Leu Ser Asp Gln Ile Gln Trp Asn Asp Val Phe Ser 435 440 445	1344
AGC CGC GCA GGT ATC CGT TAC GAC CAC ACC AAA ATG ACG CCT CAG GAA Ser Arg Ala Gly Ile Arg Tyr Asp His Thr Lys Met Thr Pro Gln Glu 450 455 460	1392
TTG AAT GCC GAC TGT CAT GCT TGT GAC AAA ACA CCG CCT GCA GCC AAC Leu Asn Ala Asp Cys His Ala Cys Asp Lys Thr Pro Pro Ala Ala Asn 465 470 475 480	1440
ACT TAT AAA GGC TGG AGC GGA TTT GTC GGC TTG GCG GCG CAG CTG AGC Thr Tyr Lys Gly Trp Ser Gly Phe Val Gly Leu Ala Ala Gln Leu Ser 485 490 495	1488
CAA ACA TGG CGT TTG GGT TAC GAT GTG ACC TCA GGT TTC CGC GTG CCG Gln Thr Trp Arg Leu Gly Tyr Asp Val Thr Ser Gly Phe Arg Val Pro 500 505 510	1536
AAT GCG TCT GAA GTG TAT TTC ACT TAC AAC CAC GGT TCG GGC ACT TGG Asn Ala Ser Glu Val Tyr Phe Thr Tyr Asn His Gly Ser Gly Thr Trp 515 520 525	1584
AAG CCT AAT CCT AAT TTG AAG GCA GAA CGC AGC ACC ACC CAC ACC CTG Lys Pro Asn Pro Asn Leu Lys Ala Glu Arg Ser Thr Thr His Thr Leu 530 535 540	1632
TCC TTG CAG GGG CGC GGC GAC AAA GGG ACA CTG GAT GCC AAC CTG TAT Ser Leu Gln Gly Arg Gly Asp Lys Gly Thr Leu Asp Ala Asn Leu Tyr 545 550 555 560	1680
CAA AGC AAT TAC CGA AAC TTC CTG TCG GAA GAG CAG AAT CTG ACT GTC Gln Ser Asn Tyr Arg Asn Phe Leu Ser Glu Glu Gln Asn Leu Thr Val 565 570 575	1728
AGC GGC ACA CCC GGC TGT ACT GAG GAG GAT GCT TAC TAC TAT AGA TGC Ser Gly Thr Pro Gly Cys Thr Glu Glu Asp Ala Tyr Tyr Arg Cys 580 585 590	1776
AGC GAC CCC TAC AAA GAA AAA CTG GAT TGG CAG ATG AAA AAT ATC GAC Ser Asp Pro Tyr Lys Glu Lys Leu Asp Trp Gln Met Lys Asn Ile Asp 595 600 605	1824
AAG GCC AGA ATC CGC GGT ATC GAG TTG ACA GGC CGT CTG AAT GTG GAC Lys Ala Arg Ile Arg Gly Ile Glu Leu Thr Gly Arg Leu Asn Val Asp 610 615 620	1872

AAA GTA GCG TCT TTT GTT CCT GAG GGT TGG AAA CTG TTC GGC TCG CTG Lys Val Ala Ser Phe Val Pro Glu Gly Trp Lys Leu Phe Gly Ser Leu 625 630 635 640	1920
GGT TAT GCG AAA AGC AAA CTG TCG GGC GAC AAC AGC CTG CTG TCC ACA Gly Tyr Ala Lys Ser Lys Leu Ser Gly Asp Asn Ser Leu Leu Ser Thr 645 650 655	1968
CAG CCG CTG AAA GTG ATT GCC GGT ATC GAC TAT GAA AGT CCG AGC GAA Gln Pro Leu Lys Val Ile Ala Gly Ile Asp Tyr Glu Ser Pro Ser Glu 660 665 670	2016
AAA TGG GGC GTA TTC TCC CGC CTG ACC TAT CTA GGC GCG AAA AAG GTC Lys Trp Gly Val Phe Ser Arg Leu Thr Tyr Leu Gly Ala Lys Lys Val 675 680 685	2064
AAA GAC GCG CAA TAC ACC GTT TAT GAA AAC AAG GGC TGG GGT ACG CCT Lys Asp Ala Gln Tyr Thr Val Tyr Glu Asn Lys Gly Trp Gly Thr Pro 690 695 700	2112
TTG CAG AAA AAG GTA AAA GAT TAC CCG TGG CTG AAC AAG TCG GCT TAT Leu Gln Lys Lys Val Lys Asp Tyr Pro Trp Leu Asn Lys Ser Ala Tyr 705 710 715 720	2160
GTG TTT GAT ATG TAC GGC TTC TAC AAA CCG GCT AAA AAC CTG ACT TTG Val Phe Asp Met Tyr Gly Phe Tyr Lys Pro Ala Lys Asn Leu Thr Leu 725 730 735	2208
CGT GCA GGC GTG TAC AAC CTG TTC AAC CGC AAA TAC ACC ACT TGG GAT Arg Ala Gly Val Tyr Asn Leu Phe Asn Arg Lys Tyr Thr Trp Asp 740 745 750	2256
TCC CTG CGC GGT TTA TAT AGC TAC AGC ACC ACC AAT GCG GTC GAC CGC Ser Leu Arg Gly Leu Tyr Ser Tyr Ser Thr Thr Asn Ala Val Asp Arg 755 760 765	2304
GAT GGC AAA GGC TTA GAC CGC TAC CGC GCC CCA GGC CGC AAT TAC GCC Asp Gly Lys Gly Leu Asp Arg Tyr Arg Ala Pro Gly Arg Asn Tyr Ala 770 775 780	2352
GTA TCG CTG GAA TGG AAG TTT TAA Val Ser Leu Glu Trp Lys Phe * 785 790	2375

## (2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 791 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Lys	Pro	Leu	Gln	Met	Pro	Pro	Ile	Ala	Ala	Leu	Leu	Gly	Ser	Ile
1				5				10						15	
Phe	Gly	Asn	Pro	Val	Phe	Ala	Ala	Asp	Glu	Ala	Ala	Thr	Glu	Thr	Thr
			20					25					30		
Pro	Val	Lys	Ala	Glu	Val	Lys	Ala	Val	Arg	Val	Lys	Gly	Gln	Arg	Asn
		35					40					45			

Ala Pro Ala Ala Val Glu Arg Val Asn Leu Asn Arg Ile Lys Gln Glu  
 50 55 60  
 Met Ile Arg Asp Asn Lys Asp Leu Val Arg Tyr Ser Thr Asp Val Gly  
 65 70 75 80  
 Leu Ser Asp Arg Ser Arg His Gln Lys Gly Phe Ala Ile Arg Gly Val  
 85 90 95  
 Glu Gly Asp Arg Val Gly Val Ser Ile Asp Gly Val Asn Leu Pro Asp  
 100 105 110  
 Ser Glu Glu Asn Ser Leu Tyr Ala Arg Tyr Gly Asn Phe Asn Ser Ser  
 115 120 125  
 Arg Leu Ser Ile Asp Pro Glu Leu Val Arg Asn Ile Asp Ile Val Lys  
 130 135 140  
 Gly Ala Asp Ser Phe Asn Thr Gly Ser Gly Ala Leu Gly Gly Gly Val  
 145 150 155 160  
 Asn Tyr Gln Thr Leu Gln Gly Arg Asp Leu Leu Leu Pro Glu Arg Gln  
 165 170 175  
 Phe Gly Val Met Met Lys Asn Gly Tyr Ser Thr Arg Asn Arg Glu Trp  
 180 185 190  
 Thr Asn Thr Leu Gly Phe Gly Val Ser Asn Asp Arg Val Asp Ala Ala  
 195 200 205  
 Leu Leu Tyr Ser Gln Arg Arg Gly His Glu Thr Glu Ser Ala Gly Lys  
 210 215 220  
 Arg Gly Tyr Pro Val Glu Gly Ala Gly Ser Gly Ala Asn Ile Arg Gly  
 225 230 235 240  
 Ser Ala Arg Gly Ile Pro Asp Pro Ser Gln His Lys Tyr His Ser Phe  
 245 250 255  
 Leu Gly Lys Ile Ala Tyr Gln Ile Asn Asp Asn His Arg Ile Gly Ala  
 260 265 270  
 Ser Leu Asn Gly Gln Gln Gly His Asn Tyr Thr Val Glu Glu Ser Tyr  
 275 280 285  
 Asn Leu Leu Ala Ser Tyr Trp Arg Glu Ala Asp Asp Val Asn Arg Arg  
 290 295 300  
 Arg Asn Thr Asn Leu Phe Tyr Glu Trp Thr Pro Glu Ser Asp Arg Leu  
 305 310 315 320  
 Ser Met Val Lys Ala Asp Val Asp Tyr Gln Lys Thr Lys Val Ser Ala  
 325 330 335  
 Val Asn Tyr Lys Gly Ser Phe Pro Thr Asn Tyr Thr Thr Trp Glu Thr  
 340 345 350  
 Glu Tyr His Lys Lys Glu Val Gly Glu Ile Tyr Asn Arg Ser Met Asp  
 355 360 365  
 Thr Thr Phe Lys Arg Ile Thr Leu Arg Met Asp Ser His Pro Leu Gln  
 370 375 380  
 Leu Gly Gly Gly Arg His Arg Leu Ser Phe Lys Thr Phe Ala Gly Gln  
 385 390 395 400

Arg Asp Phe Glu Asn Leu Asn Arg Asp Asp Tyr Tyr Phe Ser Gly Arg  
 405 410 415  
 Val Val Arg Thr Thr Asn Ser Ile Gln His Pro Val Lys Thr Thr Asn  
 420 425 430  
 Tyr Gly Phe Ser Leu Ser Asp Gln Ile Gln Trp Asn Asp Val Phe Ser  
 435 440 445  
 Ser Arg Ala Gly Ile Arg Tyr Asp His Thr Lys Met Thr Pro Gln Glu  
 450 455 460  
 Leu Asn Ala Asp Cys His Ala Cys Asp Lys Thr Pro Pro Ala Ala Asn  
 465 470 475 480  
 Thr Tyr Lys Gly Trp Ser Gly Phe Val Gly Leu Ala Ala Gln Leu Ser  
 485 490 495  
 Gln Thr Trp Arg Leu Gly Tyr Asp Val Thr Ser Gly Phe Arg Val Pro  
 500 505 510  
 Asn Ala Ser Glu Val Tyr Phe Thr Tyr Asn His Gly Ser Gly Thr Trp  
 515 520 525  
 Lys Pro Asn Pro Asn Leu Lys Ala Glu Arg Ser Thr His Thr Leu  
 530 535 540  
 Ser Leu Gln Gly Arg Gly Asp Lys Gly Thr Leu Asp Ala Asn Leu Tyr  
 545 550 555 560  
 Gln Ser Asn Tyr Arg Asn Phe Leu Ser Glu Glu Gln Asn Leu Thr Val  
 565 570 575  
 Ser Gly Thr Pro Gly Cys Thr Glu Glu Asp Ala Tyr Tyr Tyr Arg Cys  
 580 585 590  
 Ser Asp Pro Tyr Lys Glu Lys Leu Asp Trp Gln Met Lys Asn Ile Asp  
 595 600 605  
 Lys Ala Arg Ile Arg Gly Ile Glu Leu Thr Gly Arg Leu Asn Val Asp  
 610 615 620  
 Lys Val Ala Ser Phe Val Pro Glu Gly Trp Lys Leu Phe Gly Ser Leu  
 625 630 635 640  
 Gly Tyr Ala Lys Ser Lys Leu Ser Gly Asp Asn Ser Leu Leu Ser Thr  
 645 650 655  
 Gln Pro Leu Lys Val Ile Ala Gly Ile Asp Tyr Glu Ser Pro Ser Glu  
 660 665 670  
 Lys Trp Gly Val Phe Ser Arg Leu Thr Tyr Leu Gly Ala Lys Lys Val  
 675 680 685  
 Lys Asp Ala Gln Tyr Thr Val Tyr Glu Asn Lys Gly Trp Gly Thr Pro  
 690 695 700  
 Leu Gln Lys Lys Val Lys Asp Tyr Pro Trp Leu Asn Lys Ser Ala Tyr  
 705 710 715 720  
 Val Phe Asp Met Tyr Gly Phe Tyr Lys Pro Ala Lys Asn Leu Thr Leu  
 725 730 735  
 Arg Ala Gly Val Tyr Asn Leu Phe Asn Arg Lys Tyr Thr Thr Trp Asp  
 740 745 750

Ser Leu Arg Gly Leu Tyr Ser Tyr Ser Thr Thr Asn Ala Val Asp Arg  
           755                                  760                                  765

Asp Gly Lys Gly Leu Asp Arg Tyr Arg Ala Pro Gly Arg Asn Tyr Ala  
       770                                  775                                  780

Val Ser Leu Glu Trp Lys Phe  
 785                                  790

## (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2379 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..2379

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATG AAA CCA TTA CAA ATG CTC CCT ATC GCC GCG CTG GTC GGC AGT ATT	48
Met Lys Pro Leu Gln Met Leu Pro Ile Ala Ala Leu Val Gly Ser Ile	
1                                  5                                  10                                  15	
TTC GGC AAT CCG GTC TTT GCG GCA GAT GAA GCT GCA ACT GAA ACC ACA	96
Phe Gly Asn Pro Val Phe Ala Ala Asp Glu Ala Ala Thr Glu Thr Thr	
20                                  25                                  30	
CCC GTT AAG GCA GAG GTA AAA GCA GTG CGC GTT AAA GGC CAG CGC AAT	144
Pro Val Lys Ala Glu Val Lys Ala Val Arg Val Lys Gly Gln Arg Asn	
35                                  40                                  45	
GCG CCT GCG GCT GTG GAA CGC GTC AAC CTT AAC CGT ATC AAA CAA GAA	192
Ala Pro Ala Ala Val Glu Arg Val Asn Leu Asn Arg Ile Lys Gln Glu	
50                                  55                                  60	
ATG ATA CGC GAC AAC AAA GAC TTG GTG CGC TAT TCC ACC GAT GTC GGC	240
Met Ile Arg Asp Asn Lys Asp Leu Val Arg Tyr Ser Thr Asp Val Gly	
65                                  70                                  75                                  80	
TTG AGC GAC AGC GGC CGC CAT CAA AAA GGC TTT GCT GTT CGC GGC GTG	288
Leu Ser Asp Ser Gly Arg His Gln Lys Gly Phe Ala Val Arg Gly Val	
85                                  90                                  95	
GAA GGC AAC CGT GTC GGC GTG AGC ATA GAC GGC GTA AAC CTG CCT GAT	336
Glu Gly Asn Arg Val Gly Val Ser Ile Asp Gly Val Asn Leu Pro Asp	
100                                  105                                  110	
TCC GAA GAA AAC TCG CTG TAC GCC CGT TAT GGC AAC TTC AAC AGC TCG	384
Ser Glu Glu Asn Ser Leu Tyr Ala Arg Tyr Gly Asn Phe Asn Ser Ser	
115                                  120                                  125	
CGT CTG TCT ATC GAC CCC GAA CTC GTG CGC AAC ATC GAC ATC GTA AAA	432
Arg Leu Ser Ile Asp Pro Leu Val Arg Asn Ile Asp Ile Val Lys	
130                                  135                                  140	
GGG GCG GAC TCT TTC AAT ACC GGC AGC GGC GCC TTG GGC GGC GGT GTG	480
Gly Ala Asp Ser Phe Asn Thr Gly Ser Gly Ala Leu Gly Gly Gly Val	
145                                  150                                  155                                  160	



AAT TAC CAA ACC CTG CAA GGA CGT GAC TTA CTG TTG CCT GAA CGG CAG Asn Tyr Gln Thr Leu Gln Gly Arg Asp Leu Leu Leu Pro Glu Arg Gln 165 170 175	528
TTC GGC GTG ATG ATG AAA AAC GGT TAC AGC ACG CGT AAC CGT GAA TGG Phe Gly Val Met Met Lys Asn Gly Tyr Ser Thr Arg Asn Arg Glu Trp 180 185 190	576
ACA AAT ACC CTC GGT TTC GGC GTG AGC AAC GAC CGC GTG GAT GCC GCT Thr Asn Thr Leu Gly Phe Gly Val Ser Asn Asp Arg Val Asp Ala Ala 195 200 205	624
TTG CTG TAT TCG CAA CGG CGC GGC CAT GAA ACT GAA AGC GCG GGC AAG Leu Leu Tyr Ser Gln Arg Arg Gly His Glu Thr Glu Ser Ala Gly Lys 210 215 220	672
CGT GGT TAT CCG GTA GAG GGT GCT GGT AGC GGA GCG AAT ATC CGT GGT Arg Gly Tyr Pro Val Glu Gly Ala Gly Ser Gly Ala Asn Ile Arg Gly 225 230 235 240	720
TCT GCG CGC GGT ATT CCT GAT CCG TCC CAA CAC AAA TAC CAC AGC TTC Ser Ala Arg Gly Ile Pro Asp Pro Ser Gln His Lys Tyr His Ser Phe 245 250 255	768
TTG GGT AAG ATT GCT TAT CAA ATC AAC GAC AAC CAC CGC ATC GGC GCA Leu Gly Lys Ile Ala Tyr Gln Ile Asn Asp Asn His Arg Ile Gly Ala 260 265 270	816
TCG CTC AAC GGT CAG CAG GGG CAT AAT TAC ACG GTT GAA GAG TCT TAC Ser Leu Asn Gly Gln Gln Gly His Asn Tyr Thr Val Glu Glu Ser Tyr 275 280 285	864
AAC CTG CTT GCT TCT TAT TGG CGT GAA GCT GAC GAT GTC AAC AGA CGG Asn Leu Leu Ala Ser Tyr Trp Arg Glu Ala Asp Asp Val Asn Arg Arg 290 295 300	912
CGT AAC ACC AAC CTC TTT TAC GAA TGG ACG CCG GAA TCC GAC CGG TTG Arg Asn Thr Asn Leu Phe Tyr Glu Trp Thr Pro Glu Ser Asp Arg Leu 305 310 315 320	960
TCT ATG GTA AAA GCG GAT GTC GAT TAT CAA AAA ACC AAA GTA TCT GCG Ser Met Val Lys Ala Asp Val Asp Tyr Gln Lys Thr Lys Val Ser Ala 325 330 335	1008
GTC AAC TAC AAA GGT TCG TTC CCG ATA GAG GAT TCT TCC ACC TTG ACA Val Asn Tyr Lys Gly Ser Phe Pro Ile Glu Asp Ser Ser Thr Leu Thr 340 345 350	1056
CGT AAC TAC AAT CAA AAG GAC TTG GAT GAA ATC TAC AAC CGC AGT ATG Arg Asn Tyr Asn Gln Lys Asp Leu Asp Glu Ile Tyr Asn Arg Ser Met 355 360 365	1104
GAT ACC CGC TTC AAA CGC ATT ACC CTG CGT TTG GAC AGC CAT CCG TTG Asp Thr Arg Phe Lys Arg Ile Thr Leu Arg Leu Asp Ser His Pro Leu 370 375 380	1152
CAA CTC GGG GGG GGG CGA CAC CGC CTG TCG TTT AAA ACT TTC GCC AGC Gln Leu Gly Gly Gly Arg His Arg Leu Ser Phe Lys Thr Phe Ala Ser 385 390 395 400	1200
CGC CGT GAT TTT GAA AAC CTA AAC CGC GAC GAT TAT TAC TTC AGC GGC Arg Arg Asp Phe Glu Asn Leu Asn Arg Asp Asp Tyr Tyr Phe Ser Gly 405 410 415	1248
CGT GTT GTT CGA ACC ACC AGC AGT ATC CAG CAT CCG GTG AAA ACC ACC Arg Val Val Arg Thr Thr Ser Ser Ile Gln His Pro Val Lys Thr Thr 420 425 430	1296

AAC TAC GGT TTC TCA CTG TCT GAC CAA ATT CAA TGG AAC GAC GTG TTC Asn Tyr Gly Phe Ser Leu Ser Asp Gln Ile Gln Trp Asn Asp Val Phe 435 440 445	1344
AGT AGC CGC GCA GGT ATC CGT TAC GAT CAT ACC AAA ATG ACG CCT CAG Ser Ser Arg Ala Gly Ile Arg Tyr Asp His Thr Lys Met Thr Pro Gln 450 455 460	1392
GAA TTG AAT GCC GAG TGT CAT GCT TGT GAC AAA ACA CCG CCT GCA GCC Glu Leu Asn Ala Glu Cys His Ala Cys Asp Lys Thr Pro Pro Ala Ala 465 470 475 480	1440
AAC ACT TAT AAA GGC TGG AGC GGT TTT GTC GGC TTG GCG GCG CAA CTG Asn Thr Tyr Lys Gly Trp Ser Gly Phe Val Gly Leu Ala Ala Gln Leu 485 490 495	1488
AAT CAG GCT TGG CGT GTC GGT TAC GAC ATT ACT TCC GGC TAC CGT GTC Asn Gln Ala Trp Arg Val Gly Tyr Asp Ile Thr Ser Gly Tyr Arg Val 500 505 510	1536
CCC AAT GCG TCC GAA GTG TAT TTC ACT TAC AAC CAC GGT TCG GGT AAT Pro Asn Ala Ser Glu Val Tyr Phe Thr Tyr Asn His Gly Ser Gly Asn 515 520 525	1584
TGG CTG CCC AAT CCC AAC CTG AAA GCC GAG CGC ACG ACC ACC CAC ACC Trp Leu Pro Asn Pro Asn Leu Lys Ala Glu Arg Thr Thr Thr His Thr 530 535 540	1632
CTC TCT CTG CAA GGC CGC AGC GAA AAA GGT ACT TTG GAT GCC AAC CTG Leu Ser Leu Gln Gly Arg Ser Glu Lys Gly Thr Leu Asp Ala Asn Leu 545 550 555 560	1680
TAT CAA AGC AAT TAC CGC AAT TTC CTG TCT GAA GAG CAG AAG CTG ACC Tyr Gln Ser Asn Tyr Arg Asn Phe Leu Ser Glu Glu Gln Lys Leu Thr 565 570 575	1728
ACC AGC GGC GAT GTC AGC TGT ACT CAG ATG AAT TAC TAC TAC GGT ATG Thr Ser Gly Asp Val Ser Cys Thr Gln Met Asn Tyr Tyr Tyr Gly Met 580 585 590	1776
TGT AGC AAT CCT TAT TCC GAA AAA CTG GAA TGG CAG ATG CAA AAT ATC Cys Ser Asn Pro Tyr Ser Glu Lys Leu Glu Trp Gln Met Gln Asn Ile 595 600 605	1824
GAC AAG GCC AGA ATC CGC GGT ATC GAG CTG ACG GGC CGT CTG AAT GTG Asp Lys Ala Arg Ile Arg Gly Ile Glu Leu Thr Gly Arg Leu Asn Val 610 615 620	1872
GAC AAA GTA GCG TCT TTT GTT CCT GAG GGC TGG AAA CTG TTC GGC TCG Asp Lys Val Ala Ser Phe Val Pro Glu Gly Trp Lys Leu Phe Gly Ser 625 630 635 640	1920
CTG GGT TAT GCG AAA AGC AAA CTG TCG GGC GAC AAC AGC CTG CTG TCC Leu Gly Tyr Ala Lys Ser Lys Leu Ser Gly Asp Asn Ser Leu Leu Ser 645 650 655	1968
ACC CAG CCG TTG AAA GTG ATT GCC GGT ATC GAC TAT GAA AGT CCG AGC Thr Gln Pro Leu Lys Val Ile Ala Gly Ile Asp Tyr Glu Ser Pro Ser 660 665 670	2016
GAA AAA TGG GGC GTG TTC TCC CGC CTG ACC TAT CTG GGC GCG AAA AAG Glu Lys Trp Gly Val Phe Ser Arg Leu Thr Tyr Leu Gly Ala Lys Lys 675 680 685	2064
GTC AAA GAC GCG CAA TAC ACC GTT TAT GAA AAC AAG GGC TGG GGT ACG Val Lys Asp Ala Gln Tyr Thr Val Tyr Glu Asn Lys Gly Trp Gly Thr 690 695 700	2112

CCT TTG CAG AAA AAG GTA AAA GAT TAC CCG TGG CTG AAC AAG TCG GCT	2160
Pro Leu Gln Lys Lys Val Lys Asp Tyr Pro Trp Leu Asn Lys Ser Ala	
705 710 715 720	
TAT GTG TTC GAT ATG TAC GGC TTC TAC AAA CCG GTG AAA AAC CTG ACT	2208
Tyr Val Phe Asp Met Tyr Gly Phe Tyr Lys Pro Val Lys Asn Leu Thr	
725 730 735	
TTG CGT GCA GGC GTA TAT AAT GTG TTC AAC CGC AAA TAC ACC ACT TGG	2256
Leu Arg Ala Gly Val Tyr Asn Val Phe Asn Arg Lys Tyr Thr Thr Trp	
740 745 750	
GAT TCC CTG CGC GGC CTG TAT AGC TAC AGC ACC ACC AAC TCG GTC GAC	2304
Asp Ser Leu Arg Gly Leu Tyr Ser Tyr Ser Thr Thr Asn Ser Val Asp	
755 760 765	
CGC GAT GGC AAA GGC TTA GAC CGC TAC CGC GCC CCA AGC CGT AAT TAC	2352
Arg Asp Gly Lys Gly Leu Asp Arg Tyr Arg Ala Pro Ser Arg Asn Tyr	
770 775 780	
GCC GTA TCG CTG GAA TGG AAG TTT TAA	2379
Ala Val Ser Leu Glu Trp Lys Phe *	
785 790	

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 792 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Lys Pro Leu Gln Met Leu Pro Ile Ala Ala Leu Val Gly Ser Ile	
1 5 10 15	
Phe Gly Asn Pro Val Phe Ala Ala Asp Glu Ala Ala Thr Glu Thr Thr	
20 25 30	
Pro Val Lys Ala Glu Val Lys Ala Val Arg Val Lys Gly Gln Arg Asn	
35 40 45	
Ala Pro Ala Ala Val Glu Arg Val Asn Leu Asn Arg Ile Lys Gln Glu	
50 55 60	
Met Ile Arg Asp Asn Lys Asp Leu Val Arg Tyr Ser Thr Asp Val Gly	
65 70 75 80	
Leu Ser Asp Ser Gly Arg His Gln Lys Gly Phe Ala Val Arg Gly Val	
85 90 95	
Glu Gly Asn Arg Val Gly Val Ser Ile Asp Gly Val Asn Leu Pro Asp	
100 105 110	
Ser Glu Glu Asn Ser Leu Tyr Ala Arg Tyr Gly Asn Phe Asn Ser Ser	
115 120 125	
Arg Leu Ser Ile Asp Pro Glu Leu Val Arg Asn Ile Asp Ile Val Lys	
130 135 140	
Gly Ala Asp Ser Phe Asn Thr Gly Ser Gly Ala Leu Gly Gly Gly Val	
145 150 155 160	

Asn Tyr Gln Thr Leu Gln Gly Arg Asp Leu Leu Leu Pro Glu Arg Gln  
 165 170 175  
 Phe Gly Val Met Met Lys Asn Gly Tyr Ser Thr Arg Asn Arg Glu Trp  
 180 185 190  
 Thr Asn Thr Leu Gly Phe Gly Val Ser Asn Asp Arg Val Asp Ala Ala  
 195 200 205  
 Leu Leu Tyr Ser Gln Arg Arg Gly His Glu Thr Glu Ser Ala Gly Lys  
 210 215 220  
 Arg Gly Tyr Pro Val Glu Gly Ala Gly Ser Gly Ala Asn Ile Arg Gly  
 225 230 235 240  
 Ser Ala Arg Gly Ile Pro Asp Pro Ser Gln His Lys Tyr His Ser Phe  
 245 250 255  
 Leu Gly Lys Ile Ala Tyr Gln Ile Asn Asp Asn His Arg Ile Gly Ala  
 260 265 270  
 Ser Leu Asn Gly Gln Gln Gly His Asn Tyr Thr Val Glu Glu Ser Tyr  
 275 280 285  
 Asn Leu Leu Ala Ser Tyr Trp Arg Glu Ala Asp Asp Val Asn Arg Arg  
 290 295 300  
 Arg Asn Thr Asn Leu Phe Tyr Glu Trp Thr Pro Glu Ser Asp Arg Leu  
 305 310 315 320  
 Ser Met Val Lys Ala Asp Val Asp Tyr Gln Lys Thr Lys Val Ser Ala  
 325 330 335  
 Val Asn Tyr Lys Gly Ser Phe Pro Ile Glu Asp Ser Ser Thr Leu Thr  
 340 345 350  
 Arg Asn Tyr Asn Gln Lys Asp Leu Asp Glu Ile Tyr Asn Arg Ser Met  
 355 360 365  
 Asp Thr Arg Phe Lys Arg Ile Thr Leu Arg Leu Asp Ser His Pro Leu  
 370 375 380  
 Gln Leu Gly Gly Gly Arg His Arg Leu Ser Phe Lys Thr Phe Ala Ser  
 385 390 395 400  
 Arg Arg Asp Phe Glu Asn Leu Asn Arg Asp Asp Tyr Tyr Phe Ser Gly  
 405 410 415  
 Arg Val Val Arg Thr Thr Ser Ser Ile Gln His Pro Val Lys Thr Thr  
 420 425 430  
 Asn Tyr Gly Phe Ser Leu Ser Asp Gln Ile Gln Trp Asn Asp Val Phe  
 435 440 445  
 Ser Ser Arg Ala Gly Ile Arg Tyr Asp His Thr Lys Met Thr Pro Gln  
 450 455 460  
 Glu Leu Asn Ala Glu Cys His Ala Cys Asp Lys Thr Pro Pro Ala Ala  
 465 470 475 480  
 Asn Thr Tyr Lys Gly Trp Ser Gly Phe Val Gly Leu Ala Ala Gln Leu  
 485 490 495  
 Asn Gln Ala Trp Arg Val Gly Tyr Asp Ile Thr Ser Gly Tyr Arg Val  
 500 505 510

Pro Asn Ala Ser Glu Val Tyr Phe Thr Tyr Asn His Gly Ser Gly Asn  
 515 520 525  
 Trp Leu Pro Asn Pro Asn Leu Lys Ala Glu Arg Thr Thr Thr His Thr  
 530 535 540  
 Leu Ser Leu Gln Gly Arg Ser Glu Lys Gly Thr Leu Asp Ala Asn Leu  
 545 550 555 560  
 Tyr Gln Ser Asn Tyr Arg Asn Phe Leu Ser Glu Glu Gln Lys Leu Thr  
 565 570 575  
 Thr Ser Gly Asp Val Ser Cys Thr Gln Met Asn Tyr Tyr Tyr Gly Met  
 580 585 590  
 Cys Ser Asn Pro Tyr Ser Glu Lys Leu Glu Trp Gln Met Gln Asn Ile  
 595 600 605  
 Asp Lys Ala Arg Ile Arg Gly Ile Glu Leu Thr Gly Arg Leu Asn Val  
 610 615 620  
 Asp Lys Val Ala Ser Phe Val Pro Glu Gly Trp Lys Leu Phe Gly Ser  
 625 630 635 640  
 Leu Gly Tyr Ala Lys Ser Lys Leu Ser Gly Asp Asn Ser Leu Leu Ser  
 645 650 655  
 Thr Gln Pro Leu Lys Val Ile Ala Gly Ile Asp Tyr Glu Ser Pro Ser  
 660 665 670  
 Glu Lys Trp Gly Val Phe Ser Arg Leu Thr Tyr Leu Gly Ala Lys Lys  
 675 680 685  
 Val Lys Asp Ala Gln Tyr Thr Val Tyr Glu Asn Lys Gly Trp Gly Thr  
 690 695 700  
 Pro Leu Gln Lys Lys Val Lys Asp Tyr Pro Trp Leu Asn Lys Ser Ala  
 705 710 715 720  
 Tyr Val Phe Asp Met Tyr Gly Phe Tyr Lys Pro Val Lys Asn Leu Thr  
 725 730 735  
 Leu Arg Ala Gly Val Tyr Asn Val Phe Asn Arg Lys Tyr Thr Thr Trp  
 740 745 750  
 Asp Ser Leu Arg Gly Leu Tyr Ser Tyr Ser Thr Thr Asn Ser Val Asp  
 755 760 765  
 Arg Asp Gly Lys Gly Leu Asp Arg Tyr Arg Ala Pro Ser Arg Asn Tyr  
 770 775 780  
 Ala Val Ser Leu Glu Trp Lys Phe  
 785 790

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2378 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: CDS

(B) LOCATION: 1..2373

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATG	AAA	CCA	TTA	CAC	ATG	CTT	CCT	ATT	GCC	GCG	CTG	GTC	GGC	AGT	ATT	48
Met	Lys	Pro	Leu	His	Met	Leu	Pro	Ile	Ala	Ala	Leu	Val	Gly	Ser	Ile	
1				5					10					15		
TTC	GGC	AAT	CCG	GTC	TTG	GCA	GCG	GAT	GAA	GCT	GCA	ACC	GAA	ACC	ACA	96
Phe	Gly	Asn	Pro	Val	Leu	Ala	Ala	Asp	Glu	Ala	Ala	Thr	Glu	Thr	Thr	
			20					25					30			
CCC	GTT	AAA	GCA	GAG	ATA	AAA	GAA	GTG	CGC	GTT	AAA	GAC	CAG	CTT	AAT	144
Pro	Val	Lys	Ala	Glu	Ile	Lys	Glu	Val	Arg	Val	Lys	Asp	Gln	Leu	Asn	
		35					40					45				
GCG	CCT	GCA	ACC	GTG	GAA	CGT	GTC	AAC	CTC	GGC	CGC	ATT	CAA	CAG	GAA	192
Ala	Pro	Ala	Thr	Val	Glu	Arg	Val	Asn	Leu	Gly	Arg	Ile	Gln	Gln	Glu	
	50					55				60						
ATG	ATA	CGC	GAC	AAC	AAA	GAC	TTG	GTG	CGT	TAC	TCC	ACC	GAC	GTC	GGC	240
Met	Ile	Arg	Asp	Asn	Lys	Asp	Leu	Val	Arg	Tyr	Ser	Thr	Asp	Val	Gly	
65				70						75					80	
TTG	AGC	GAT	AGC	GGC	CGC	CAT	CAA	AAA	GGC	TTT	GCT	GTG	CGC	GGC	GTG	288
Leu	Ser	Asp	Ser	Gly	Arg	His	Gln	Lys	Gly	Phe	Ala	Val	Arg	Gly	Val	
				85					90					95		
GAA	GGC	AAC	CGT	GTC	GGT	GTC	AGC	ATT	GAC	GGC	GTG	AGC	CTG	CCT	GAT	336
Glu	Gly	Asn	Arg	Val	Gly	Val	Ser	Ile	Asp	Gly	Val	Ser	Leu	Pro	Asp	
			100					105					110			
TCG	GAA	GAA	AAC	TCA	CTG	TAT	GCA	CGT	TAT	GGC	AAC	TTC	AAC	AGC	TCG	384
Ser	Glu	Glu	Asn	Ser	Leu	Tyr	Ala	Arg	Tyr	Gly	Asn	Phe	Asn	Ser	Ser	
		115					120					125				
CGC	CTG	TCT	ATC	GAC	CCC	GAA	CTC	GTG	CGC	AAC	ATC	GAA	ATC	GCG	AAG	432
Arg	Leu	Ser	Ile	Asp	Pro	Glu	Leu	Val	Arg	Asn	Ile	Glu	Ile	Ala	Lys	
	130					135					140					
GGC	GCT	GAC	TCT	TTC	AAT	ACC	GGT	AGC	GGC	GCA	TTG	GGT	GGC	GGC	GTG	480
Gly	Ala	Asp	Ser	Phe	Asn	Thr	Gly	Ser	Gly	Ala	Leu	Gly	Gly	Gly	Val	
145					150					155					160	
AAT	TAC	CAA	ACC	CTG	CAA	GGA	CAT	GAT	TTG	CTG	TTG	GAC	GAC	AGG	CAA	528
Asn	Tyr	Gln	Thr	Leu	Gln	Gly	His	Asp	Leu	Leu	Leu	Asp	Asp	Arg	Gln	
			165						170					175		
TTC	GGC	GTG	ATG	ATG	AAA	AAC	GGT	TAC	AGC	AGC	CGC	AAC	CGC	GAA	TGG	576
Phe	Gly	Val	Met	Met	Lys	Asn	Gly	Tyr	Ser	Ser	Arg	Asn	Arg	Glu	Trp	
			180					185					190			
ACA	AAT	ACA	CTC	GGT	TTC	GGT	GTG	AGC	AAC	GAC	CGC	GTG	GAT	GCC	GCT	624
Thr	Asn	Thr	Leu	Gly	Phe	Gly	Val	Ser	Asn	Asp	Arg	Val	Asp	Ala	Ala	
		195					200					205				
TTG	CTG	TAT	TCG	CAA	CGT	CGC	GGT	CAT	GAG	ACC	GAA	AGC	GCG	GGC	GAG	672
Leu	Leu	Tyr	Ser	Gln	Arg	Arg	Gly	His	Glu	Thr		Ser	Ala	Gly	Glu	
	210					215					220					
CGT	GGC	TAT	CCG	GTA	GAG	GGT	GCT	GGC	AGC	GGA	GCA	ATT	ATC	CGT	GGT	720
Arg	Gly	Tyr	Pro	Val	Glu	Gly	Ala	Gly	Ser	Gly	Ala	Ile	Ile	Arg	Gly	
225					230					235					240	
TCG	TCA	CGC	GGT	ATC	CCT	GAT	CCG	TCC	AAA	CAC	AAA	TAC	CAC	AAC	TTC	768
Ser	Ser	Arg	Gly	Ile	Pro	Asp	Pro	Ser	Lys	His	Lys	Tyr	His	Asn	Phe	
				245					250					255		

TTG GGT AAG ATT GCT TAT CAA ATC AAC GAC AAG CAC CGC ATC GGC CCA Leu Gly Lys Ile Ala Tyr Gln Ile Asn Asp Lys His Arg Ile Gly Pro	816
260 265 270	
TCG TTT AAC GGC CAG CAG GGG CAT AAT TAC ACG ATT GAA GAG TCT TAT Ser Phe Asn Gly Gln Gln Gly His Asn Tyr Thr Ile Glu Glu Ser Tyr	864
275 280 285	
AAC CTG ACC GCT TCT TCC TGG CGC GAA GCC GAT GAC GTA AAC AGA CGG Asn Leu Thr Ala Ser Ser Trp Arg Glu Ala Asp Asp Val Asn Arg Arg	912
290 295 300	
CGC AAT GCC AAC CTC TTT TAC GAA TGG ACG CCT GAT TCA AAT TGG CTG Arg Asn Ala Asn Leu Phe Tyr Glu Trp Thr Pro Asp Ser Asn Trp Leu	960
305 310 315 320	
TCG TCT TTG AAG GCG GAC TTC GAT TAT CAG ACA ACC AAA GTG GCG GCG Ser Ser Leu Lys Ala Asp Phe Asp Tyr Gln Thr Thr Lys Val Ala Ala	1008
325 330 335	
GTT AAC AAC AAA GGC TCG TTC CCG ACG GAT TAT TCC ACC TGG ACG CGC Val Asn Asn Lys Gly Ser Phe Pro Thr Asp Tyr Ser Thr Trp Thr Arg	1056
340 345 350	
AAC TAT AAT CAG AAG GAT TTG GAG AAT ATA TAC AAC CGC AGC ATG GAC Asn Tyr Asn Gln Lys Asp Leu Glu Asn Ile Tyr Asn Arg Ser Met Asp	1104
355 360 365	
ACC CGA TTC AAA CGT TTT ACT TTG CGT ATG GAC AGC CAA CCG TTG CAA Thr Arg Phe Lys Arg Phe Thr Leu Arg Met Asp Ser Gln Pro Leu Gln	1152
370 375 380	
CTG GGC GGC CAA CAT CGC TTG TCG CTT AAA ACT TTC GCC AGT CGG CGT Leu Gly Gly Gln His Arg Leu Ser Leu Lys Thr Phe Ala Ser Arg Arg	1200
385 390 395 400	
GAG TTT GAA AAC TTA AAC CGC GAC GAT TAT TAC TTC AGC GAA AGA GTA Glu Phe Glu Asn Leu Asn Arg Asp Asp Tyr Tyr Phe Ser Glu Arg Val	1248
405 410 415	
TCC CGT ACT ACC AGC TCG ATT CAA CAC CCC GTG AAA ACC ACT AAT TAT Ser Arg Thr Thr Ser Ser Ile Gln His Pro Val Lys Thr Thr Asn Tyr	1296
420 425 430	
GGT TTC TCA CTG TCT GAT CAA ATC CAA TGG AAC GAC GTG TTC AGC AGC Gly Phe Ser Leu Ser Asp Gln Ile Gln Trp Asn Asp Val Phe Ser Ser	1344
435 440 445	
CGT GCA GAT ATC CGT TAC GAT CAT ACC AAA ATG ACG CCT CAG GAA TTG Arg Ala Asp Ile Arg Tyr Asp His Thr Lys Met Thr Pro Gln Glu Leu	1392
450 455 460	
AAT GCC GAG TGT CAT GCT TGT GAC AAA ACA CCG CCT GCA GCC AAT ACT Asn Ala Glu Cys His Ala Cys Asp Lys Thr Pro Pro Ala Ala Asn Thr	1440
465 470 475 480	
TAT AAA GGC TGG AGC GGA TTT GTC GGT TTG GCG GCG CAA CTG AAT CAG Tyr Lys Gly Trp Ser Gly Phe Val Gly Leu Ala Ala Gln Leu Asn Gln	1488
485 490 495	
GCT TGG CAT GTC GGT TAC GAC ATT ACT TCC GGC TAC CGT GTC CCC AAT Ala Trp His Val Gly Tyr Asp Ile Thr Ser Gly Tyr Arg Val Pro Asn	1536
500 505 510	
GCG TCC GAA GTG TAT TTC ACT TAC AAC CAC GGT TCG GGT AAT TGG CTG Ala Ser Glu Val Tyr Phe Thr Tyr Asn His Gly Ser Gly Asn Trp Leu	1584
515 520 525	

CCC Pro 530	AAT Asn	CCC Pro	AAC Asn	CTG Leu	AAA Lys	GCC Ala 535	GAG Glu	CGC Arg	AGC Ser	ACC Thr	ACC Thr 540	CAC His	ACC Thr	CTG Leu	TCT Ser	1632
CTG Leu 545	CAA Gln	GGC Gly	CGC Arg	AGC Ser	GAA Glu 550	AAA Lys	GGT Gly	ACT Thr	TTG Leu	GAT Asp 555	GCC Ala	AAC Asn	CTG Leu	TAT Tyr	CAA Gln 560	1680
AAC Asn	AAT Asn	TAC Tyr	CGC Arg	AAC Asn 565	TTC Phe	TTG Leu	TCT Ser	GAA Glu	GAG Glu 570	CAG Gln	AAG Lys	CTG Leu	ACC Thr	ACC Thr 575	AGC Ser	1728
GGC Gly	GAT Asp	GTC Val	GGC Gly 580	TGT Cys	ACT Thr	CAG Gln	ATG Met	AAT Asn 585	TAC Tyr	TAC Tyr	TAC Tyr	GGT Gly	ATG Met 590	TGT Cys	AGC Ser	1776
AAT Asn	CCT Pro	TAT Tyr 595	TCC Ser	GAA Glu	AAA Lys	CCG Pro	GAA Glu 600	TGG Trp	CAG Gln	ATG Met	CAA Gln	AAT Asn 605	ATC Ile	GAT Asp	AAG Lys	1824
GCC Ala 610	CGA Arg	ATC Ile	CGT Arg	GGT Gly	CTT Leu	GAG Glu 615	CTG Leu	ACA Thr	GGC Gly	CGT Arg	CTG Leu 620	AAT Asn	GTG Val	ACA Thr	AAA Lys	1872
GTA Val 625	GCG Ala	TCT Ser	TTT Phe	GTT Val	CCT Pro 630	GAG Glu	GGC Gly	TGG Trp	AAA Lys	TTG Leu 635	TTC Phe	GGC Gly	TCG Ser	CTG Leu	GGT Gly 640	1920
TAT Tyr	GCG Ala	AAA Lys	AGC Ser	AAA Lys 645	CTG Leu	TCG Ser	GGC Gly	GAC Asp	AAC Asn 650	AGC Ser	CTG Leu	CTG Leu	TCC Ser	ACA Thr 655	CAG Gln	1968
CCG Pro	CCG Pro	AAA Lys	GTG Val 660	ATT Ile	GCC Ala	GGT Gly	GTC Val	GAC Asp 665	TAC Tyr	GAA Glu	AGC Ser	CCG Pro	AGC Ser 670	GAA Glu	AAA Lys	2016
TGG Trp	GGT Gly	GTG Val 675	TTC Phe	TCC Ser	CGC Arg	CTG Leu	ACT Thr 680	TAT Tyr	CTG Leu	GGT Gly	GCG Ala	AAA Lys 685	AAG Lys	GCC Ala	AAA Lys	2064
GAC Asp 690	GCG Ala	CAA Gln	TAC Tyr	ACC Thr	GTT Val	TAT Tyr 695	GAA Glu	AAC Asn	AAG Lys	GGC Gly	CGG Arg	GGT Gly	ACG Thr	CCT Pro	TTG Leu	2112
CAG Gln 705	AAA Lys	AAG Lys	GTA Val	AAA Lys	GAT Asp 710	TAC Tyr	CCG Pro	TGG Trp	CTG Leu	AAC Asn 715	AAG Lys	TCG Ser	GCT Ala	TAT Tyr	GTG Val 720	2160
TTT Phe	GAT Asp	ATG Met	TAC Tyr	GGC Gly 725	TTC Phe	TAC Tyr	AAA Lys	CTG Leu	GCT Ala 730	AAA Lys	AAC Asn	CTG Leu	ACT Thr	TTG Leu 735	CGT Arg	2208
GCA Ala	GGC Gly	GTA Val	TAT Tyr	AAT Asn	GTG Val	TTC Phe	AAC Asn	CGC Arg 745	AAA Lys	TAC Tyr	ACC Thr	ACT Thr	TGG Trp 750	GAT Asp	TCC Ser	2256
CTG Leu	CGC Arg	GGT Gly 755	TTG Leu	TAT Tyr	AGC Ser	TAC Tyr	AGC Ser	ACC Thr 760	ACC Thr	AAC Asn	GCG Ala	GTC Val 765	GAC Asp	CGA Arg	GAT Asp	2304
GGC Gly 770	AAA Lys	GGC Gly	TTA Leu	GAC Asp	CGC Arg	TAC Tyr 775	CGC Arg	GCC Ala	TCA Ser	GGC Gly	CGT Asn 780	AAT Asn	TAC Tyr	GCC Ala	GTA Val	2352
TCG Ser 785	CTG Leu	GAT Asp	TGG Trp	AAG Lys	TTT Phe 790	TGA	ATTCC									2378



## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 790 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

Met Lys Pro Leu His Met Leu Pro Ile Ala Ala Leu Val Gly Ser Ile
 1           5           10           15
Phe Gly Asn Pro Val Leu Ala Ala Asp Glu Ala Ala Thr Glu Thr Thr
          20           25           30
Pro Val Lys Ala Glu Ile Lys Glu Val Arg Val Lys Asp Gln Leu Asn
          35           40           45
Ala Pro Ala Thr Val Glu Arg Val Asn Leu Gly Arg Ile Gln Gln Glu
          50           55           60
Met Ile Arg Asp Asn Lys Asp Leu Val Arg Tyr Ser Thr Asp Val Gly
          65           70           75           80
Leu Ser Asp Ser Gly Arg His Gln Lys Gly Phe Ala Val Arg Gly Val
          85           90           95
Glu Gly Asn Arg Val Gly Val Ser Ile Asp Gly Val Ser Leu Pro Asp
          100          105          110
Ser Glu Glu Asn Ser Leu Tyr Ala Arg Tyr Gly Asn Phe Asn Ser Ser
          115          120          125
Arg Leu Ser Ile Asp Pro Glu Leu Val Arg Asn Ile Glu Ile Ala Lys
          130          135          140
Gly Ala Asp Ser Phe Asn Thr Gly Ser Gly Ala Leu Gly Gly Gly Val
          145          150          155          160
Asn Tyr Gln Thr Leu Gln Gly His Asp Leu Leu Leu Asp Asp Arg Gln
          165          170          175
Phe Gly Val Met Met Lys Asn Gly Tyr Ser Ser Arg Asn Arg Glu Trp
          180          185          190
Thr Asn Thr Leu Gly Phe Gly Val Ser Asn Asp Arg Val Asp Ala Ala
          195          200          205
Leu Leu Tyr Ser Gln Arg Arg Gly His Glu Thr Glu Ser Ala Gly Glu
          210          215          220
Arg Gly Tyr Pro Val Glu Gly Ala Gly Ser Gly Ala Ile Ile Arg Gly
          225          230          235          240
Ser Ser Arg Gly Ile Pro Asp Pro Ser Lys His Lys Tyr His Asn Phe
          245          250          255
Leu Gly Lys Ile Ala Tyr Gln Ile Asn Asp Lys His Arg Ile Gly Pro
          260          265          270
Ser Phe Asn Gly Gln Gln Gly His Asn Tyr Thr Ile Glu Glu Ser Tyr
          275          280          285
Asn Leu Thr Ala Ser Ser Trp Arg Glu Ala Asp Asp Val Asn Arg Arg
          290          295          300

```

Arg Asn Ala Asn Leu Phe Tyr Glu Trp Thr Pro Asp Ser Asn Trp Leu  
 305 310 315 320  
 Ser Ser Leu Lys Ala Asp Phe Asp Tyr Gln Thr Thr Lys Val Ala Ala  
 325 330 335  
 Val Asn Asn Lys Gly Ser Phe Pro Thr Asp Tyr Ser Thr Trp Thr Arg  
 340 345 350  
 Asn Tyr Asn Gln Lys Asp Leu Glu Asn Ile Tyr Asn Arg Ser Met Asp  
 355 360 365  
 Thr Arg Phe Lys Arg Phe Thr Leu Arg Met Asp Ser Gln Pro Leu Gln  
 370 375 380  
 Leu Gly Gly Gln His Arg Leu Ser Leu Lys Thr Phe Ala Ser Arg Arg  
 385 390 395 400  
 Glu Phe Glu Asn Leu Asn Arg Asp Asp Tyr Tyr Phe Ser Glu Arg Val  
 405 410 415  
 Ser Arg Thr Thr Ser Ser Ile Gln His Pro Val Lys Thr Thr Asn Tyr  
 420 425 430  
 Gly Phe Ser Leu Ser Asp Gln Ile Gln Trp Asn Asp Val Phe Ser Ser  
 435 440 445  
 Arg Ala Asp Ile Arg Tyr Asp His Thr Lys Met Thr Pro Gln Glu Leu  
 450 455 460  
 Asn Ala Glu Cys His Ala Cys Asp Lys Thr Pro Pro Ala Ala Asn Thr  
 465 470 475 480  
 Tyr Lys Gly Trp Ser Gly Phe Val Gly Leu Ala Ala Gln Leu Asn Gln  
 485 490 495  
 Ala Trp His Val Gly Tyr Asp Ile Thr Ser Gly Tyr Arg Val Pro Asn  
 500 505 510  
 Ala Ser Glu Val Tyr Phe Thr Tyr Asn His Gly Ser Gly Asn Trp Leu  
 515 520 525  
 Pro Asn Pro Asn Leu Lys Ala Glu Arg Ser Thr Thr His Thr Leu Ser  
 530 535 540  
 Leu Gln Gly Arg Ser Glu Lys Gly Thr Leu Asp Ala Asn Leu Tyr Gln  
 545 550 555 560  
 Asn Asn Tyr Arg Asn Phe Leu Ser Glu Glu Gln Lys Leu Thr Thr Ser  
 565 570 575  
 Gly Asp Val Gly Cys Thr Gln Met Asn Tyr Tyr Tyr Gly Met Cys Ser  
 580 585 590  
 Asn Pro Tyr Ser Glu Lys Pro Glu Trp Gln Met Gln Asn Ile Asp Lys  
 595 600 605  
 Ala Arg Ile Arg Gly Leu Glu Leu Thr Gly Arg Leu Asn Val Thr Lys  
 610 615 620  
 Val Ala Ser Phe Val Pro Glu Gly Trp Lys Leu Phe Gly Ser Leu Gly  
 625 630 635 640  
 Tyr Ala Lys Ser Lys Leu Ser Gly Asp Asn Ser Leu Leu Ser Thr Gln  
 645 650 655

Pro Pro Lys Val Ile Ala Gly Val Asp Tyr Glu Ser Pro Ser Glu Lys  
                     660                    665                    670  
 Trp Gly Val Phe Ser Arg Leu Thr Tyr Leu Gly Ala Lys Lys Ala Lys  
                     675                    680                    685  
 Asp Ala Gln Tyr Thr Val Tyr Glu Asn Lys Gly Arg Gly Thr Pro Leu  
                     690                    695                    700  
 Gln Lys Lys Val Lys Asp Tyr Pro Trp Leu Asn Lys Ser Ala Tyr Val  
                     705                    710                    715                    720  
 Phe Asp Met Tyr Gly Phe Tyr Lys Leu Ala Lys Asn Leu Thr Leu Arg  
                     725                    730                    735  
 Ala Gly Val Tyr Asn Val Phe Asn Arg Lys Tyr Thr Thr Trp Asp Ser  
                     740                    745                    750  
 Leu Arg Gly Leu Tyr Ser Tyr Ser Thr Thr Asn Ala Val Asp Arg Asp  
                     755                    760                    765  
 Gly Lys Gly Leu Asp Arg Tyr Arg Ala Ser Gly Arg Asn Tyr Ala Val  
                     770                    775                    780  
 Ser Leu Asp Trp Lys Phe  
                     785                    790

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 641 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Gln Gln Gln His Leu Phe Arg Leu Asn Ile Leu Cys Leu Ser Leu  
 1                    5                    10                    15  
 Met Thr Ala Leu Pro Val Tyr Ala Glu Asn Val Gln Ala Glu Gln Ala  
                     20                    25                    30  
 Gln Glu Lys Gln Leu Asp Thr Ile Val Lys Ala Lys Lys Gln Lys Thr  
                     35                    40                    45  
 Arg Arg Asp Asn Glu Val Thr Gly Leu Gly Lys Leu Val Lys Ser Ser  
                     50                    55                    60  
 Asp Thr Leu Ser Lys Glu Gln Val Leu Asn Ile Arg Asp Leu Thr Arg  
                     65                    70                    75                    80  
 Tyr Asp Pro Gly Ile Ala Val Val Glu Gln Gly Arg Gly Ala Ser Ser  
                     85                    90                    95  
 Gly Tyr Ser Ile Arg Gly Met Asp Lys Asn Arg Val Ser Leu Thr Val  
                     100                    105                    110  
 Asp Gly Val Ser Gln Ile Gln Ser Tyr Thr Ala Gln Ala Ala Leu Gly  
                     115                    120                    125  
 Gly Thr Arg Thr Ala Gly Ser Ser Gly Ala Ile Asn Glu Ile Glu Tyr  
                     130                    135                    140

Glu Asn Val Lys Ala Val Glu Ile Ser Lys Gly Ser Asn Ser Ser Glu  
 145 150 155 160  
 Tyr Gly Asn Gly Ala Leu Ala Gly Ser Val Ala Phe Gln Thr Lys Thr  
 165 170 175  
 Ala Ala Asp Ile Ile Gly Glu Gly Lys Gln Trp Gly Ile Gln Ser Lys  
 180 185 190  
 Thr Ala Tyr Ser Gly Lys Asp His Ala Leu Thr Gln Ser Leu Ala Leu  
 195 200 205  
 Ala Gly Arg Ser Gly Gly Ala Glu Ala Leu Leu Ile Tyr Thr Lys Arg  
 210 215 220  
 Arg Gly Arg Glu Ile His Ala His Lys Asp Ala Gly Lys Gly Val Gln  
 225 230 235 240  
 Ser Phe Asn Arg Leu Pro Ile Cys Arg Phe Gly Asn Asn Thr Tyr Thr  
 245 250 255  
 Asp Cys Thr Pro Arg Asn Ile Gly Gly Asn Gly Tyr Tyr Ala Val  
 260 265 270  
 Gln Asp Asn Val Arg Leu Gly Arg Trp Ala Asp Val Gly Ala Gly Ile  
 275 280 285  
 Arg Tyr Asp Tyr Arg Ser Thr His Ser Glu Asp Lys Ser Val Ser Thr  
 290 295 300  
 Gly Thr His Arg Asn Leu Ser Trp Asn Ala Gly Val Val Leu Lys Pro  
 305 310 315 320  
 Phe Thr Trp Met Asp Leu Thr Tyr Arg Ala Ser Thr Gly Phe Arg Leu  
 325 330 335  
 Pro Ser Phe Ala Glu Met Tyr Gly Trp Arg Ala Gly Glu Ser Leu Lys  
 340 345 350  
 Thr Leu Asp Leu Lys Pro Glu Lys Ser Phe Asn Arg Glu Ala Gly Ile  
 355 360 365  
 Val Phe Lys Gly Asp Phe Gly Asn Leu Glu Ala Ser Tyr Phe Asn Asn  
 370 375 380  
 Ala Tyr Arg Asp Leu Ile Ala Phe Gly Tyr Glu Thr Arg Thr Gln Asn  
 385 390 395 400  
 Gly Gln Thr Ser Ala Ser Gly Asp Pro Gly Tyr Arg Asn Ala Gln Asn  
 405 410 415  
 Ala Arg Ile Ala Gly Ile Asn Ile Leu Gly Lys Ile Asp Trp His Gly  
 420 425 430  
 Val Trp Gly Gly Leu Pro Asp Gly Leu Tyr Ser Thr Leu Ala Tyr Asn  
 435 440 445  
 Arg Ile Lys Val Lys Asp Ala Asp Arg Ala Asp Arg Thr Phe Val Thr  
 450 455 460  
 Ser Tyr Leu Phe Asp Ala Val Gln Pro Ser Arg Tyr Val Leu Gly Leu  
 465 470 475 480  
 Gly Tyr Asp His Pro Asp Gly Ile Trp Gly Ile Asn Thr Met Phe Thr  
 485 490 495  
 Tyr Ser Lys Ala Lys Ser Val Asp Glu Leu Leu Gly Ser Gln Ala Leu

```

          500                      505                      510
Leu Asn Gly Asn Ala Asn Ala Lys Lys Ala Ala Ser Arg Arg Thr Arg
   515                      520                      525
Pro Trp Tyr Val Thr Asp Val Ser Gly Tyr Tyr Asn Ile Lys Lys His
   530                      535                      540
Leu Thr Leu Arg Ala Gly Val Tyr Asn Leu Leu Asn Tyr Arg Tyr Val
   545                      550                      555
Thr Trp Glu Asn Val Arg Gln Thr Ala Gly Gly Ala Val Asn Gln His
   565                      570                      575
Lys Asn Val Gly Val Tyr Asn Arg Tyr Ala Ala Pro Gly Arg Asn Tyr
   580                      585                      590
Thr Phe Ser Leu Glu Met Lys Phe
   595                      600

```

## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 607 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

Met Asn Lys Lys His Gly Phe Gln Leu Thr Leu Thr Ala Leu Ala Val
 1          5          10          15
Ala Ala Ala Phe Pro Ser Tyr Ala Ala Asn Pro Glu Thr Ala Ala Pro
 20          25          30
Asp Ala Ala Gln Thr Gln Ser Leu Lys Glu Val Thr Val Arg Ala Ala
 35          40          45
Lys Val Gly Arg Arg Ser Lys Glu Ala Thr Gly Leu Gly Lys Ile Ala
 50          55          60
Lys Thr Ser Glu Thr Leu Asn Lys Glu Gln Val Leu Gly Ile Arg Asp
 65          70          75          80
Leu Thr Arg Tyr Asp Pro Gly Val Ala Val Val Glu Gln Gly Asn Gly
 85          90          95

Ala Ser Gly Gly Tyr Ser Ile Arg Gly Val Asp Lys Asn Arg Val Ala
100          105          110
Val Ser Val Asp Gly Val Ala Gln Ile Gln Ala Phe Thr Val Gln Gly
115          120          125
Ser Leu Ser Gly Tyr Gly Gly Arg Gly Gly Ser Gly Ala Ile Asn Glu
130          135          140
Ile Glu Tyr Glu Asn Ile Ser Thr Val Glu Ile Asp Lys Gly Ala Gly
145          150          155          160
Ser Ser Asp His Gly Ser Gly Ala Leu Gly Gly Ala Val Ala Phe Arg
165          170          175

```

Thr Lys Glu Ala Ala Asp Leu Ile Ser Asp Gly Lys Ser Trp Gly Ile  
 180 185 190  
 Gln Ala Lys Thr Ala Tyr Gly Ser Lys Asn Arg Gln Phe Met Lys Ser  
 195 200 205  
 Leu Gly Ala Gly Phe Ser Lys Asp Gly Trp Glu Gly Leu Leu Ile Arg  
 210 215 220  
 Thr Glu Arg Gln Gly Arg Glu Thr His Pro His Gly Asp Ile Ala Asp  
 225 230 235 240  
 Gly Val Ala Tyr Gly Ile Asn Arg Leu Ser Val Cys Gly Tyr Ile Glu  
 245 250 255  
 Thr Leu Arg Ser Arg Lys Cys Val Pro Arg Lys Ile Asn Gly Ser Asn  
 260 265 270  
 Ile His Ile Ser Leu Asn Asp Arg Phe Ser Ile Gly Lys Tyr Phe Asp  
 275 280 285  
 Phe Ser Leu Gly Gly Arg Tyr Asp Arg Lys Asn Phe Thr Thr Ser Glu  
 290 295 300  
 Glu Leu Val Arg Ser Gly Arg Tyr Val Asp Arg Ser Trp Asn Ser Gly  
 305 310 315 320  
 Ile Val Phe Lys Pro Asn Arg His Phe Ser Leu Ser Tyr Arg Ala Ser  
 325 330 335  
 Ser Gly Phe Arg Thr Pro Ser Phe Gln Glu Leu Phe Gly Ile Asp Ile  
 340 345 350  
 Tyr His Asp Tyr Pro Lys Gly Trp Gln Arg Pro Ala Leu Lys Ser Glu  
 355 360 365  
 Lys Ala Ala Asn Arg Glu Ile Gly Leu Gln Trp Lys Gly Asp Phe Gly  
 370 375 380  
 Phe Leu Glu Ile Ser Ser Phe Arg Asn Arg Tyr Thr Asp Met Ile Ala  
 385 390 395 400  
 Val Ala Asp His Lys Thr Lys Leu Pro Asn Gln Ala Gly Gln Leu Thr  
 405 410 415  
 Glu Ile Asp Ile Arg Asp Tyr Tyr Asn Ala Gln Asn Met Ser Leu Gln  
 420 425 430  
 Gly Val Asn Ile Leu Gly Lys Ile Asp Trp Asn Gly Val Tyr Gly Lys  
 435 440 445  
 Leu Pro Glu Gly Leu Tyr Thr Thr Leu Ala Tyr Asn Arg Ile Lys Pro  
 450 455 460  
 Lys Ser Val Ser Asn Arg Pro Gly Leu Ser Leu Arg Ser Tyr Ala Leu  
 465 470 475 480  
 Asp Ala Val Gln Pro Ser Arg Tyr Val Leu Gly Phe Gly Tyr Asp Gln  
 485 490 495  
 Pro Glu Gly Lys Trp Gly Ala Asn Ile Met Leu Thr Tyr Ser Lys Gly  
 500 505 510  
 Lys Asn Pro Asp Glu Leu Ala Tyr Leu Ala Gly Asp Gln Lys Arg Tyr  
 515 520 525

```

Ser Thr Lys Arg Ala Ser Ser Ser Trp Ser Thr Ala Asp Val Ser Ala
530                               535                               540

Tyr Leu Asn Leu Lys Lys Arg Leu Thr Leu Arg Ala Ala Ile Tyr Asn
545                               550                               555                               560

Ile Gly Asn Tyr Arg Tyr Val Thr Trp Glu Ser Leu Arg Gln Thr Ala
565                               570                               575

Glu Ser Thr Ala Asn Arg His Gly Gly Asp Ser Asn Tyr Gly Arg Tyr
580                               585                               590

Ala Ala Pro Gly Arg Asn Phe Ser Leu Ala Leu Glu Met Lys Phe
595                               600                               605

```

## (2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AAACAGGTCT CGGCATAG

18

## (2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGCGAATTCA AACAGGTCTC GGCATAG

27

## (2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 33 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGCGAATTCA AAAACTTCCA TTCCAGCGAT ACG

33

## (2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TAAAACTTCC ATTCCAGCGA TACG

24

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AAACAGGTCT CGGCATAG

18

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CGCGAATTCA AACAGGTCTC GGCATAG

27

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 33 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGCGAATTCA AAAACTTCCA TTCCAGCGAT ACG

33

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid



(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TAAAACTTCC ATTCCAGCGA TACG

24

**WHAT WE CLAIM IS:**

1. An isolated and purified recombinant nucleic acid encoding a hemoglobin receptor protein from a *Neisseria* species.
2. An isolated and purified recombinant nucleic acid according to Claim 1, wherein the nucleic acid encodes a hemoglobin receptor protein having an amino acid sequence that is the amino acid sequence depicted as Seq. I.D. No. 2.
3. An isolated and purified recombinant nucleic acid according to Claim 1, wherein the nucleic acid encodes a hemoglobin receptor protein having an amino acid sequence that is the amino acid sequence depicted as Seq. I.D. No. 4.
4. An isolated and purified recombinant nucleic acid according to Claim 1, wherein the nucleic acid encodes a hemoglobin receptor protein having an amino acid sequence that is the amino acid sequence depicted as Seq. I.D. No. 6.
5. An isolated and purified recombinant nucleic acid according to Claim 1, wherein the nucleic acid encodes a hemoglobin receptor protein having an amino acid sequence that is the amino acid sequence depicted as Seq. I.D. No. 8.
6. A homogeneous preparation of a hemoglobin receptor protein from a *Neisseria* species.
7. The hemoglobin receptor protein of Claim 6, wherein the protein has an amino acid sequence that is the amino acid sequence depicted as Seq. I.D. No. 2.
8. The hemoglobin receptor protein of Claim 6, wherein the protein has an amino acid sequence that is the amino acid sequence depicted as Seq. I.D. No. 4.
9. The hemoglobin receptor protein of Claim 6, wherein the protein has an amino acid sequence that is the amino acid sequence depicted as Seq. I.D. No. 6.
10. The hemoglobin receptor protein of Claim 6, wherein the protein has an amino acid sequence that is the amino acid sequence depicted as Seq. I.D. No. 8.
11. A recombinant expression construct comprising a nucleic acid that encodes a hemoglobin receptor protein from a *Neisseria* species.

12. A transformed cell culture comprising the recombinant expression construct of Claim 11.
13. A recombinant expression construct according to Claim 11, wherein the nucleic acid encodes a hemoglobin receptor protein having an amino acid sequence that is the amino acid sequence depicted as Seq. I.D. No. 2.
14. A recombinant expression construct according to Claim 11, wherein the nucleic acid encodes a hemoglobin receptor protein having an amino acid sequence that is the amino acid sequence depicted as Seq. I.D. No. 4.
15. A recombinant expression construct according to Claim 11, wherein the nucleic acid encodes a hemoglobin receptor protein having an amino acid sequence that is the amino acid sequence depicted as Seq. I.D. No. 6.
16. A recombinant expression construct according to Claim 11, wherein the nucleic acid encodes a hemoglobin receptor protein having an amino acid sequence that is the amino acid sequence depicted as Seq. I.D. No. 8.
17. A transformed cell culture comprising the recombinant expression construct of Claims 13, 14, 15 or 16.
18. An antibody or antigen-binding fragment thereof that is immunologically reactive with an antigenic epitope of a hemoglobin receptor protein from a *Neisseria* species.
19. An antibody according to Claim 18 that is a monoclonal antibody.
20. An antibody or antigen-binding fragment thereof according to Claim 18 that is immunologically reactive with an antigenic epitope of the hemoglobin receptor protein depicted as Seq. I.D. No. 2.
21. An antibody or antigen-binding fragment thereof according to Claim 18 that is immunologically reactive with an antigenic epitope of the hemoglobin receptor protein depicted as Seq. I.D. No. 4.
22. An antibody or antigen-binding fragment thereof according to Claim 18 that is immunologically reactive with an antigenic epitope of the hemoglobin receptor protein depicted as Seq. I.D. No. 6.
23. An antibody or antigen-binding fragment thereof according to Claim 18 that is immunologically reactive with an antigenic epitope of the hemoglobin receptor protein depicted as Seq. I.D. No. 8.

24. An antigenic epitope of a hemoglobin receptor protein from a *Neisseria* species.

25. The antigenic epitope of Claim 24 wherein the hemoglobin receptor protein is the protein depicted as Seq. I.D. No. 2.

5 26. The antigenic epitope of Claim 24 wherein the hemoglobin receptor protein is the protein depicted as Seq. I.D. No. 4.

27. The antigenic epitope of Claim 24 wherein the hemoglobin receptor protein is the protein depicted as Seq. I.D. No. 6.

10 28. The antigenic epitope of Claim 24 wherein the hemoglobin receptor protein is the protein depicted as Seq. I.D. No. 8.

29. A diagnostic reagent for diagnosing a disease state in a human, wherein the disease state is caused by bacteria of a *Neisseria* species, the diagnostic reagent comprising an antibody according to Claims 18, 20, 21, 22, or 23.

15 30. A diagnostic reagent for diagnosing a disease state in a human, wherein the disease state is caused by bacteria of a *Neisseria* species, the diagnostic reagent comprising an antibody according to Claim 19.

31. A diagnostic reagent for diagnosing a disease state in a human, wherein the disease state is caused by bacteria of a *Neisseria* species, the diagnostic reagent comprising the nucleic acid of Claim 1.

20 32. A diagnostic reagent for diagnosing a disease state in a human, wherein the disease state is caused by bacteria of a *Neisseria* species, the diagnostic reagent comprising the nucleic acid of Claims 2, 3, 4 or 5.

25 33. A therapeutic agent for treating a disease state in a human, wherein the disease state is caused by bacteria of a *Neisseria* species, the therapeutic agent comprising an antibody according to Claim 18, 20, 21, 22, or 23.

34. A therapeutic agent for treating a disease state in a human, wherein the disease state is caused by bacteria of a *Neisseria* species, the therapeutic agent comprising an antibody according to Claim 19.

30 35. A therapeutic agent for treating a disease state in a human, wherein the disease state is caused by bacteria of a *Neisseria* species, the therapeutic agent comprising the nucleic acid of Claim 1 or antisense homologue thereof.

36. A therapeutic agent for treating a disease state in a human, wherein the disease state is caused by bacteria of a *Neisseria* species, the therapeutic agent comprising the nucleic acid of Claims 2, 3, 4, or 5 or antisense homologue thereof.

37. A therapeutic agent for treating a disease state in a human, wherein the disease state is caused by bacteria of a *Neisseria* species, the therapeutic agent comprising the recombinant expression construct of Claims 11, 13, 14, 15 or 16 or a homologue thereof that expresses the nucleic acid encoding a hemoglobin receptor in an antisense orientation.

38. An antibody according to Claims 20, 21, 22 or 23 that is a monoclonal antibody.

39. A cell line that produces the monoclonal antibody of Claims 19 or 38.

40. A method of treating a disease in a human caused by bacteria of a *Neisseria* species, the method comprising the step of administering a therapeutically-effective amount of the therapeutic agent of Claims 33, 34, 35, 36, or 37 in a pharmaceutically-acceptable carrier.

41. A method of diagnosing a disease in a human caused by bacteria of a *Neisseria* species, the method comprising the steps of contacting an amount of a detectably-labeled diagnostic reagent of Claims 29, 30, 31, or 32 to a biological sample from the human under conditions wherein the diagnostic reagent specifically binds to the *Neisseria* bacteria and detecting an amount of the specific binding to the biological sample.

42. A vaccine that is effective in providing immunization against infection of a human with a bacteria of *Neisseria* species comprising a hemoglobin binding protein or antigenic fragment thereof.

43. The vaccine of Claim 42 comprising the hemoglobin receptor protein of Claims 6, 7, 8, 9, or 10.

44. The vaccine of Claim 42 comprising a nucleic acid encoding a hemoglobin receptor protein from a *Neisseria* species or antigenic fragment thereof.

45. A vaccine according to Claim 44 comprising the nucleic acid of Claims 2, 3, 4, 5, 11, 13, 14, 15, or 16.

46. The vaccine of Claim 42 comprising cells of the transformed cell culture of Claim 17.

47. A vaccine according to Claim 46 wherein the cells are attenuated bacterial cells.

48. A vaccine according to Claim 47 wherein the cells are *Salmonella* cells.

5           49. The vaccine of Claim 42 comprising the epitope of the hemoglobin receptor protein of Claims 24, 25, 26, 27 or 28.

1/22

Fig. 1

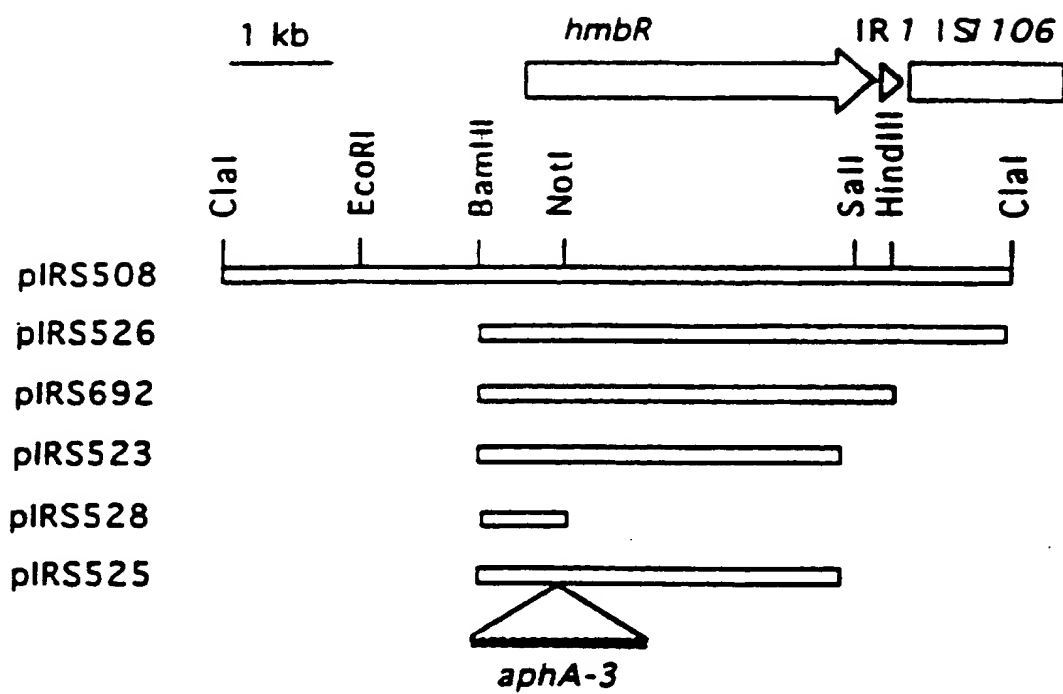


Fig. 2

BNSDOCID: <WO 9612020A2 | >



Fig. 2 (cont'd.)

BNSDOCID: &lt;WO 9612020A2 | &gt;

Fig. 3

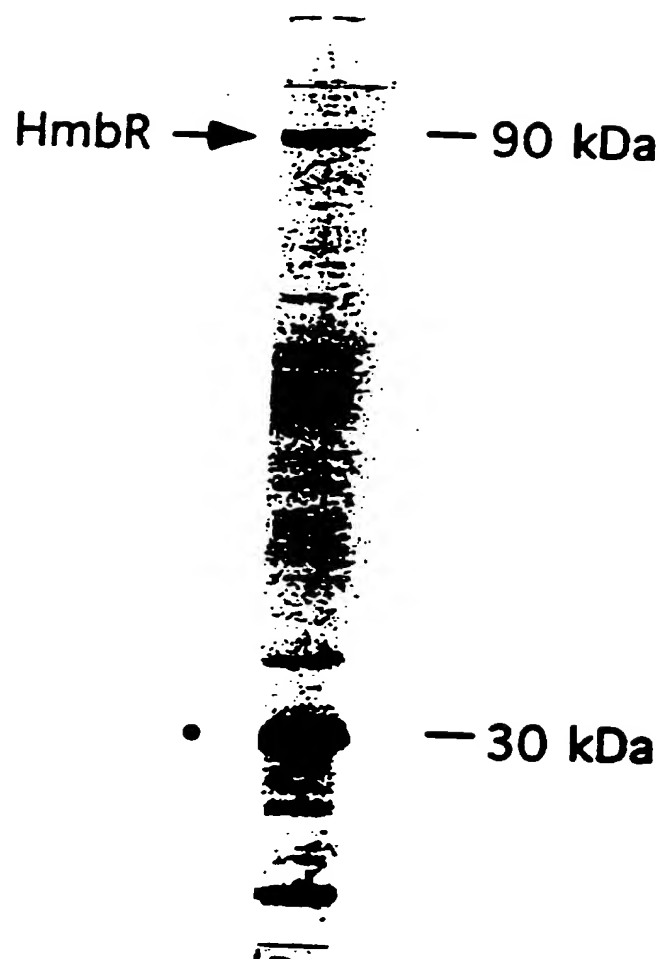


Fig. 4

LBPA	MNKGNGFQLTALAVAAAFPSYAANPETAAPDAAGTVS	30
HQBR	HQPLQHLPIAALVGSIFGN-PVFAADEAATETETKAE-----VKAVRV	43
	.. .. .	
TBPIH	KTRRINVTGLGKLVKSEDTLKKQVLNIRDLTRYDFGIAVVEQGRGAS	97
LBPA	-GRUSKEATGLGKIATSETLQKEQVLGIRDLTRYDFGVAVVEQNGASC	99
HQBR	KQQRMA-PAAVERV--KLNRLKQEMIRDKDLVRYSTOVGLESDSGRQK-	89
	.. .. .	
TBPIH	GYSIRGMDNRVSLTVQGVQIQSYTAQAALGGTRTAGSSGAINETIETN	147
LBPA	GYSIRGVDNRVAVSVQGVQIQAFVQGSLSGGYGGGGSGAINETIETN	149
HQBR	GFAVRGVGNRVGVSIDGVNLPDS--ETNSLYARYGNFNSRLS-IDPEL	136
	.. .. .	
TBPIH	VKAVEISXGNSSEYNGALAGSVAFOTKTAADIIGEGKQNGIQSKTAYS	197
LBPA	ISTVEIDKGAGSSDHGSGALGGAVAFRTKEAADLISDGKSWGIOAKTAYG	199
HQBR	VRNIDIVKGADSFNTGSGALGGGVNYOTLOGRDLLLPERQFGVHGGNGYS	186
	.. .. .	
TBPIH	GKDHALTQSLALAGRSGGAELLYTKRGREIHAKKDAKGVO-SFNRL	246
LBPA	SKNRQFQKSLGAGFSKDGWEGLLIRTERQGRETHPHGDIADGVAYGINRL	249
HQBR	TRNRDWTNLGFGVSNDRVDAALLYSORGHETESAG-----	223
	.. .. .	
TBPIH	PICRFGNNTYT-DCTPRNIGNGYAAVQDNVRLGRWADVGAGIRYDYS	601
LBPA	SVCGYIETLASRKCVPRKINGSNIHISLNDRFSIGKYTDFSLGGRYDRUN	635
HQBR	-----SSIQHPVKTTNYGFSLSDOIQNDVFSRAGIRYDMTK	460
	.. .. .	
TBPIH	THSED-----KSVSTGTGRLSWNAGVVLKP--FTWMDLTYRSTGF	641
LBPA	PTTSE-----ELVREGRYVDRLSWNSGIVFKP--NRHFLSYRASSGF	675
HQBR	MTPOELNAECHACDKTPPAANTYKWSGFGVGLAAQLNQAMRVGYDITSGY	510
	.. .. .	
TBPIH	RLPSFAEMYGWRA---GESLKTLDLKPEKSFNREAGIVFKGDFGNLEAS	687
LBPA	RTPSFQELFGIDYNDYPKGWQRPALKSEKAANREIGLQWKGDFGFLEIS	725
HQBR	RVPNASEVY-FTYNGSGMWLPHNPKAERTTHTLSLQGRSEKGTLDAN	559
	.. .. .	
TBPIH	YFNAYRDLIAFGYET---RTQNGOTSASGDFGYR-----	719
LBPA	SFRNRYTIMIAVADHKTELPHQACOLTEIDIRDY-----	760
HQBR	LYQSNRYNLS---EEQKLT-SDGVSTQNGYNYGHCNPFYSKLEMQH	605
	.. .. .	
TBPIH	-NAQNARIAGINILGKIDMGVMGGLPDG--LYSTLAYNRKVKQADIRA	766
LBPA	-NAQNSLQGVNILGKIDMGVYGLPEG--LYTTLAYNRKPKSVENRP	807
HQBR	QNIDKARIRGIELTGRINVDKVASFVPEGMILFGSLGYAKSKLSC----	650
	.. .. .	
TBPIH	DRTFVTSYLPDAVQPSRYVLGLQYENFDGIMGINTWQTYSKAJSVDE---	813
LBPA	GLSL-RSYALDAVQPSRYVLQGYDQFEGKMGANIKLTYSGKQYFDE---	853
HQBR	DNLSLLST-----QPLKVLAGIDYESPSERMGVFSRLTYLGAJKVKDAGY	694
	.. .. .	
TBPIH	-LLGSOALLNGMANAKKAASRTTRWYVTVDSGYNYIKKGLTLRAGVYHL	862
LBPA	-L----ATLAGDQR-KYSTKRASSMSTADVSAYLNLKRLTLRAAIYNI	897
HQBR	TVYENKAGCTFLQKVEDYPMNLKLSAYVVDNYGPFYKPVQGLTLRAGVYHV	744
	.. .. .	
TBPIH	LNRYRYVTEDVRC--TAGGAVNQDNVGVYNYRYAAPGRNNTTFSLEKSF	908
LBPA	GNRYRYVTESLRQ--TAESTANRUGGDSNYGRYAAPGRNPELALSKF	943
HQBR	PNRYTTHDSLAGLYSTSTNSVDKGGGLDRYTRAPSRNYAVELEKSF	792
	.. .. .	

Fig. 5

1 2 3 4

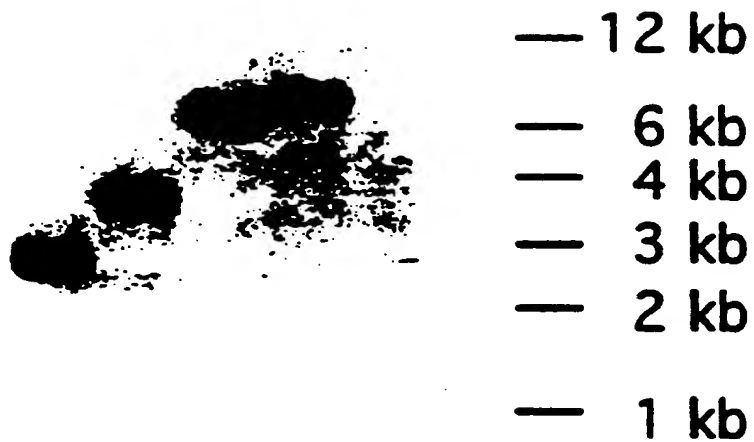
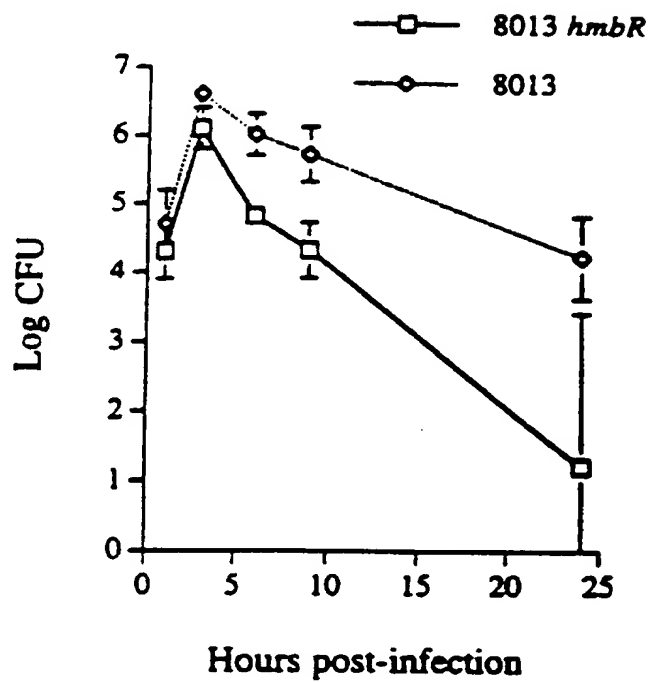


Fig. 6



**Figure 7**

ATG AAA CCA TTA CAA ATG CCC CCT ATC GCC GCG CTG CTC GGC AGT ATT	48
Met Lys Pro Leu Gln Met Pro Pro Ile Ala Ala Leu Leu Gly Ser Ile	
1 5 10 15	
TTC GGC AAT CCG GTC TTT GCG GCA GAT GAA GCT GCA ACT GAA ACC ACA	96
Phe Gly Asn Pro Val Phe Ala Ala Asp Glu Ala Ala Thr Glu Thr Thr	
20 25 30	
CCC GTT AAG GCA GAG GTA AAA GCA GTG CGC GTT AAA GGT CAG CGC AAT	144
Pro Val Lys Ala Glu Val Lys Ala Val Arg Val Lys Gly Gln Arg Asn	
35 40 45	
GCG CCT GCG GCT GTG GAA CGC GTC AAC CTT AAC CGT ATC AAA CAA GAA	192
Ala Pro Ala Ala Val Glu Arg Val Asn Leu Asn Arg Ile Lys Gln Glu	
50 55 60	
ATG ATA CGC GAC AAT AAA GAC TTG GTG CGC TAT TCC ACC GAT GTC GGC	240
Met Ile Arg Asp Asn Lys Asp Leu Val Arg Tyr Ser Thr Asp Val Gly	
65 70 75 80	
TTG AGC GAC AGG AGC CGT CAT CAA AAA GGC TTT GCC ATT CGC GGC GTG	288
Leu Ser Asp Arg Ser Arg His Gln Lys Gly Phe Ala Ile Arg Gly Val	
85 90 95	
GAA GGC GAC CGT GTC GGC GTT AGT ATT GAC GGC GTA AAC CTG CCT GAT	336
Glu Gly Asp Arg Val Gly Val Ser Ile Asp Gly Val Asn Leu Pro Asp	
100 105 110	
TCC GAA GAA AAC TCG CTG TAC GCC CGT TAT GGC AAC TTC AAC AGC TCG	384
Ser Glu Glu Asn Ser Leu Tyr Ala Arg Tyr Gly Asn Phe Asn Ser Ser	
115 120 125	
CGT CTG TCT ATC GAC CCC GAA CTC GTG CGC AAC ATC GAC ATC GTA AAA	432
Arg Leu Ser Ile Asp Pro Glu Leu Val Arg Asn Ile Asp Ile Val Lys	
130 135 140	
GGG GCG GAC TCT TTC AAT ACC GGC AGC GGC GCC TTG GGC GGC GGT GTG	480
Gly Ala Asp Ser Phe Asn Thr Gly Ser Gly Ala Leu Gly Gly Gly Val	
145 150 155 160	
AAT TAC CAA ACC CTG CAA GGA CGT GAC TTA CTG TTG CCT GAA CGG CAG	528
Asn Tyr Gln Thr Leu Gln Gly Arg Asp Leu Leu Leu Pro Glu Arg Gln	
165 170 175	
TTC GGC GTG ATG ATG AAA AAC GGT TAC AGC ACG CGT AAC CGT GAA TGG	576
Phe Gly Val Met Met Lys Asn Gly Tyr Ser Thr Arg Asn Arg Glu Trp	
180 185 190	
ACA AAT ACC CTC GGT TTC GGC GTG AGC AAC GAC CGC GTG GAT GCC GCT	624
Thr Asn Thr Leu Gly Phe Gly Val Ser Asn Asp Arg Val Asp Ala Ala	
195 200 205	
TTG CTG TAT TCG CAA CGG CGC GGC CAT GAA ACT GAA AGC GCG GGC AAG	672
Leu Leu Tyr Ser Gln Arg Arg Gly His Glu Thr Glu Ser Ala Gly Lys	
210 215 220	
CGT GGT TAT CCG GTA GAG GGT GCT GGT AGC GGA GCG AAT ATC CGT GGT	720
Arg Gly Tyr Pro Val Glu Gly Ala Gly Ser Gly Ala Asn Ile Arg Gly	
225 230 235 240	
TCT GCG CGC GGT ATT CCT GAT CCG TCC CAA CAC AAA TAC CAC AGC TTC	768
Ser Ala Arg Gly Ile Pro Asp Pro Ser Gln His Lys Tyr His Ser Phe	
245 250 255	
TTG GGT AAG ATT GCT TAT CAA ATC AAC GAC AAC CAC CGC ATC GGC GCA	816
Leu Gly Lys Ile Ala Tyr Gln Ile Asn Asp Asn His Arg Ile Gly Ala	
260 265 270	

Figure 7 (cont'd.)

TCG CTC AAC GGT CAG CAG GGG CAT AAT TAC ACG GTT GAA GAG TCT TAC Ser Leu Asn Gly Gln Gln Gly His Asn Tyr Thr Val Glu Glu Ser Tyr 275 280 285	864
AAC CTG CTT GCT TCT TAT TGG CGT GAA GCT GAC GAT GTC AAC AGA CGG Asn Leu Leu Ala Ser Tyr Trp Arg Glu Ala Asp Asp Val Asn Arg Arg 290 295 300	912
CGT AAC ACC AAC CTC TTT TAC GAA TGG ACG CCG GAA TCC GAC CGG TTG Arg Asn Thr Asn Leu Phe Tyr Glu Trp Thr Pro Glu Ser Asp Arg Leu 305 310 315 320	960
TCT ATG GTA AAA GCG GAT GTC GAT TAT CAA AAA ACC AAA GTA TCT GCG Ser Met Val Lys Ala Asp Val Asp Tyr Gln Lys Thr Lys Val Ser Ala 325 330 335	1008
GTC AAC TAC AAA GGT TCG TTC CCG ACG AAT TAC ACC ACA TGG GAA ACC Val Asn Tyr Lys Gly Ser Phe Pro Thr Asn Tyr Thr Thr Trp Glu Thr 340 345 350	1056
GAG TAC CAT AAA AAG GAA GTT GGC GAA ATC TAT AAC CGC AGC ATG GAT Glu Tyr His Lys Lys Glu Val Gly Glu Ile Tyr Asn Arg Ser Met Asp 355 360 365	1104
ACA ACC TTC AAA CGT ATT ACG CTG CGT ATG GAC AGC CAT CCG TTG CAA Thr Thr Phe Lys Arg Ile Thr Leu Arg Met Asp Ser His Pro Leu Gln 370 375 380	1152
CTC GGG GGG GGG CGA CAC CGC CTG TCG TTC AAA ACC TTT GCC GGG CAG Leu Gly Gly Gly Arg His Arg Leu Ser Phe Lys Thr Phe Ala Gly Gln 385 390 395 400	1200
CGT GAT TTT GAA AAC TTA AAC CGC GAC GAT TAC TAC TTC AGC GGC CGT Arg Asp Phe Glu Asn Leu Asn Arg Asp Asp Tyr Tyr Phe Ser Gly Arg 405 410 415	1248
GTT GTT CGA ACC ACC AAC AGT ATC CAG CAT CCG GTG AAA ACC ACC AAC Val Val Arg Thr Thr Asn Ser Ile Gln His Pro Val Lys Thr Thr Asn 420 425 430	1296
TAC GGT TTC TCG CTG TCC GAC CAA ATC CAA TGG AAC GAC GTG TTC AGT Tyr Gly Phe Ser Leu Ser Asp Gln Ile Gln Trp Asn Asp Val Phe Ser 435 440 445	1344
AGC CGC GCA GGT ATC CGT TAC GAC CAC ACC AAA ATG ACG CCT CAG GAA Ser Arg Ala Gly Ile Arg Tyr Asp His Thr Lys Met Thr Pro Gln Glu 450 455 460	1392
TTG AAT GCC GAC TGT CAT GCT TGT GAC AAA ACA CCG CCT GCA GCC AAC Leu Asn Ala Asp Cys His Ala Cys Asp Lys Thr Pro Pro Ala Ala Asn 465 470 475 480	1440
ACT TAT AAA GGC TGG AGC GGA TTT GTC GGC TTG GCG GCG CAG CTG AGC Thr Tyr Lys Gly Trp Ser Gly Phe Val Gly Leu Ala Ala Gln Leu Ser 485 490 495	1488
CAA ACA TGG CGT TTG GGT TAC GAT GTG ACC TCA GGT TTC CGC GTG CCG Gln Thr Trp Arg Leu Gly Tyr Asp Val Thr Ser Gly Phe Arg Val Pro 500 505 510	1536
AAT GCG TCT GAA GTG TAT TTC ACT TAC AAC CAC GGT TCG GGC ACT TGG Asn Ala Ser Glu Val Tyr Phe Thr Tyr Asn His Gly Ser Gly Thr Trp 515 520 525	1584

Figure 7 (cont'd.)

AAG CCT AAT CCT AAT TTG AAG GCA GAA CGC AGC ACC ACC CAC ACC CTG Lys Pro Asn Pro Asn Leu Lys Ala Glu Arg Ser Thr Thr His Thr Leu 530 535 540	1632
TCC TTG CAG GGG CGC GGC GAC AAA GGG ACA CTG GAT GCC AAC CTG TAT Ser Leu Gln Gly Arg Gly Asp Lys Gly Thr Leu Asp Ala Asn Leu Tyr 545 550 555 560	1680
CAA AGC AAT TAC CGA AAC TTC CTG TCG GAA GAG CAG AAT CTG ACT GTC Gln Ser Asn Tyr Arg Asn Phe Leu Ser Glu Glu Gln Asn Leu Thr Val 565 570 575	1728
AGC GGC ACA CCC GGC TGT ACT GAG GAG GAT GCT TAC TAC TAT AGA TGC Ser Gly Thr Pro Gly Cys Thr Glu Glu Asp Ala Tyr Tyr Arg Cys 580 585 590	1776
AGC GAC CCC TAC AAA GAA AAA CTG GAT TGG CAG ATG AAA AAT ATC GAC Ser Asp Pro Tyr Lys Glu Lys Leu Asp Trp Gln Met Lys Asn Ile Asp 595 600 605	1824
AAG GCC AGA ATC CGC GGT ATC GAG TTG ACA GGC CGT CTG AAT GTG GAC Lys Ala Arg Ile Arg Gly Ile Glu Leu Thr Gly Arg Leu Asn Val Asp 610 615 620	1872
AAA GTA GCG TCT TTT GTT CCT GAG GGT TGG AAA CTG TTC GGC TCG CTG Lys Val Ala Ser Phe Val Pro Glu Gly Trp Lys Leu Phe Gly Ser Leu 625 630 635 640	1920
GGT TAT GCG AAA AGC AAA CTG TCG GGC GAC AAC AGC CTG CTG TCC ACA Gly Tyr Ala Lys Ser Lys Leu Ser Gly Asp Asn Ser Leu Leu Ser Thr 645 650 655	1968
CAG CCG CTG AAA GTG ATT GCC GGT ATC GAC TAT GAA AGT CCG AGC GAA Gln Pro Leu Lys Val Ile Ala Gly Ile Asp Tyr Glu Ser Pro Ser Glu 660 665 670	2016
AAA TGG GGC GTA TTC TCC CGC CTG ACC TAT CTA GGC GCG AAA AAG GTC Lys Trp Gly Val Phe Ser Arg Leu Thr Tyr Leu Gly Ala Lys Lys Val 675 680 685	2064
AAA GAC GCG CAA TAC ACC GTT TAT GAA AAC AAG GGC TGG GGT ACG CCT Lys Asp Ala Gln Tyr Thr Val Tyr Glu Asn Lys Gly Trp Gly Thr Pro 690 695 700	2112
TTG CAG AAA AAG GTA AAA GAT TAC CCG TGG CTG AAC AAG TCG GCT TAT Leu Gln Lys Lys Val Lys Asp Tyr Pro Trp Leu Asn Lys Ser Ala Tyr 705 710 715 720	2160
GTG TTT GAT ATG TAC GGC TTC TAC AAA CCG GCT AAA AAC CTG ACT TTG Val Phe Asp Met Tyr Gly Phe Tyr Lys Pro Ala Lys Asn Leu Thr Leu 725 730 735	2208
CGT GCA GGC GTG TAC AAC CTG TTC AAC CGC AAA TAC ACC ACT TGG GAT Arg Ala Gly Val Tyr Asn Leu Phe Asn Arg Lys Tyr Thr Thr Trp Asp 740 745 750	2256
TCC CTG CGC GGT TTA TAT AGC TAC AGC ACC ACC AAT GCG GTC GAC CGC Ser Leu Arg Gly Leu Tyr Ser Tyr Ser Thr Thr Asn Ala Val Asp Arg 755 760 765	2304
GAT GGC AAA GGC TTA GAC CGC TAC CGC GCC CCA GGC CGC AAT TAC GCC Asp Gly Lys Gly Leu Asp Arg Tyr Arg Ala Pro Gly Arg Asn Tyr Ala 770 775 780	2352



Figure 7 (cont'd.)

GTA	TCG	CTG	GAA	TGG	AAG	TTT	TAA	
Val	Ser	Leu	Glu	Trp	Lys	Phe	*	
785					790			

2375

Figure 8

ATG AAA CCA TTA CAA ATG CTC CCT ATC GCC GCG CTG GTC GGC AGT ATT	48
Met Lys Pro Leu Gln Met Leu Pro Ile Ala Ala Leu Val Gly Ser Ile	
1 5 10 15	
TTC GGC AAT CCG GTC TTT GCG GCA GAT GAA GCT GCA ACT GAA ACC ACA	96
Phe Gly Asn Pro Val Phe Ala Ala Asp Glu Ala Ala Thr Glu Thr Thr	
20 25 30	
CCC GTT AAG GCA GAG GTA AAA GCA GTG CGC GTT AAA GGC CAG CGC AAT	144
Pro Val Lys Ala Glu Val Lys Ala Val Arg Val Lys Gly Gln Arg Asn	
35 40 45	
GCG CCT GCG GCT GTG GAA CGC GTC AAC CTT AAC CGT ATC AAA CAA GAA	192
Ala Pro Ala Ala Val Glu Arg Val Asn Leu Asn Arg Ile Lys Gln Glu	
50 55 60	
ATG ATA CGC GAC AAC AAA GAC TTG GTG CGC TAT TCC ACC GAT GTC GGC	240
Met Ile Arg Asp Asn Lys Asp Leu Val Arg Tyr Ser Thr Asp Val Gly	
65 70 75 80	
TTG AGC GAC AGC GGC CGC CAT CAA AAA GGC TTT GCT GTT CGC GGC GTG	288
Leu Ser Asp Ser Gly Arg His Gln Lys Gly Phe Ala Val Arg Gly Val	
85 90 95	
GAA GGC AAC CGT GTC GGC GTG AGC ATA GAC GGC GTA AAC CTG CCT GAT	336
Glu Gly Asn Arg Val Gly Val Ser Ile Asp Gly Val Asn Leu Pro Asp	
100 105 110	
TCC GAA GAA AAC TCG CTG TAC GCC CGT TAT GGC AAC TTC AAC AGC TCG	384
Ser Glu Glu Asn Ser Leu Tyr Ala Arg Tyr Gly Asn Phe Asn Ser Ser	
115 120 125	
CGT CTG TCT ATC GAC CCC GAA CTC GTG CGC AAC ATC GAC ATC GTA AAA	432
Arg Leu Ser Ile Asp Pro Glu Leu Val Arg Asn Ile Asp Ile Val Lys	
130 135 140	
GGG GCG GAC TCT TTC AAT ACC GGC AGC GGC GCC TTG GGC GGC GGT GTG	480
Gly Ala Asp Ser Phe Asn Thr Gly Ser Gly Ala Leu Gly Gly Gly Val	
145 150 155 160	
AAT TAC CAA ACC CTG CAA GGA CGT GAC TTA CTG TTG CCT GAA CGG CAG	528
Asn Tyr Gln Thr Leu Gln Gly Arg Asp Leu Leu Leu Pro Glu Arg Gln	
165 170 175	
TTC GGC GTG ATG ATG AAA AAC GGT TAC AGC ACG CGT AAC CGT GAA TGG	576
Phe Gly Val Met Met Lys Asn Gly Tyr Ser Thr Arg Asn Arg Glu Trp	
180 185 190	
ACA AAT ACC CTC GGT TTC GGC GTG AGC AAC GAC CGC GTG GAT GCC GCT	624
Thr Asn Thr Leu Gly Phe Gly Val Ser Asn Asp Arg Val Asp Ala Ala	
195 200 205	
TTG CTG TAT TCG CAA CGG CGC GGC CAT GAA ACT GAA AGC GCG GGC AAG	672
Leu Leu Tyr Ser Gln Arg Arg Gly His Glu Thr Glu Ser Ala Gly Lys	
210 215 220	
CGT GGT TAT CCG GTA GAG GGT GCT GGT AGC GGA GCG AAT ATC CGT GGT	720
Arg Gly Tyr Pro Val Glu Gly Ala Gly Ser Gly Ala Asn Ile Arg Gly	
225 230 235 240	
TCT GCG CGC GGT ATT CCT GAT CCG TCC CAA CAC AAA TAC CAC AGC TTC	768
Ser Ala Arg Gly Ile Pro Asp Pro Ser Gln His Lys Tyr His Ser Phe	
245 250 255	

Figure 8 (cont.'d).

TTG GGT AAG ATT GCT TAT CAA ATC AAC GAC AAC CAC CGC ATC GGC GCA Leu Gly Lys Ile Ala Tyr Gln Ile Asn Asp Asn His Arg Ile Gly Ala 260 265 270	816
TCG CTC AAC GGT CAG CAG GGG CAT AAT TAC ACG GTT GAA GAG TCT TAC Ser Leu Asn Gly Gln Gln Gly His Asn Tyr Thr Val Glu Glu Ser Tyr 275 280 285	864
AAC CTG CTT GCT TCT TAT TGG CGT GAA GCT GAC GAT GTC AAC AGA CGG Asn Leu Leu Ala Ser Tyr Trp Arg Glu Ala Asp Asp Val Asn Arg Arg 290 295 300	912
CGT AAC ACC AAC CTC TTT TAC GAA TGG ACG CCG GAA TCC GAC CGG TTG Arg Asn Thr Asn Leu Phe Tyr Glu Trp Thr Pro Glu Ser Asp Arg Leu 305 310 315 320	960
TCT ATG GTA AAA GCG GAT GTC GAT TAT CAA AAA ACC AAA GTA TCT GCG Ser Met Val Lys Ala Asp Val Asp Tyr Gln Lys Thr Lys Val Ser Ala 325 330 335	1008
GTC AAC TAC AAA GGT TCG TTC CCG ATA GAG GAT TCT TCC ACC TTG ACA Val Asn Tyr Lys Gly Ser Phe Pro Ile Glu Asp Ser Ser Thr Leu Thr 340 345 350	1056
CGT AAC TAC AAT CAA AAG GAC TTG GAT GAA ATC TAC AAC CGC AGT ATG Arg Asn Tyr Asn Gln Lys Asp Leu Asp Glu Ile Tyr Asn Arg Ser Met 355 360 365	1104
GAT ACC CGC TTC AAA CGC ATT ACC CTG CGT TTG GAC AGC CAT CCG TTG Asp Thr Arg Phe Lys Arg Ile Thr Leu Arg Leu Asp Ser His Pro Leu 370 375 380	1152
CAA CTC GGG GGG GGG CGA CAC CGC CTG TCG TTT AAA ACT TTC GCC AGC Gln Leu Gly Gly Gly Arg His Arg Leu Ser Phe Lys Thr Phe Ala Ser 385 390 395 400	1200
CGC CGT GAT TTT GAA AAC CTA AAC CGC GAC GAT TAT TAC TTC AGC GGC Arg Arg Asp Phe Glu Asn Leu Asn Arg Asp Asp Tyr Tyr Phe Ser Gly 405 410 415	1248
CGT GTT GTT CGA ACC ACC AGC AGT ATC CAG CAT CCG GTG AAA ACC ACC Arg Val Val Arg Thr Thr Ser Ser Ile Gln His Pro Val Lys Thr Thr 420 425 430	1296
AAC TAC GGT TTC TCA CTG TCT GAC CAA ATT CAA TGG AAC GAC GTG TTC Asn Tyr Gly Phe Ser Leu Ser Asp Gln Ile Gln Trp Asn Asp Val Phe 435 440 445	1344
AGT AGC CGC GCA GGT ATC CGT TAC GAT CAT ACC AAA ATG ACG CCT CAG Ser Ser Arg Ala Gly Ile Arg Tyr Asp His Thr Lys Met Thr Pro Gln 450 455 460	1392
GAA TTG AAT GCC GAG TGT CAT GCT TGT GAC AAA ACA CCG CCT GCA GCC Glu Leu Asn Ala Glu Cys His Ala Cys Asp Lys Thr Pro Pro Ala Ala 465 470 475 480	1440
AAC ACT TAT AAA GGC TGG AGC GGT TTT GTC GGC TTG GCG GCG CAA CTG Asn Thr Tyr Lys Gly Trp Ser Gly Phe Val Gly Leu Ala Ala Gln Leu 485 490 495	1488
AAT CAG GCT TGG CGT GTC GGT TAC GAC ATT ACT TCC GGC TAC CGT GTC Asn Gln Ala Trp Arg Val Gly Tyr Asp Ile Thr Ser Gly Tyr Arg Val 500 505 510	1536

Figure 8 (cont'd.)

CCC AAT GCG TCC GAA GTG TAT TTC ACT TAC AAC CAC GGT TCG GGT AAT Pro Asn Ala Ser Glu Val Tyr Phe Thr Tyr Asn His Gly Ser Gly Asn 515 520 525	1584
TGG CTG CCC AAT CCC AAC CTG AAA GCC GAG CGC ACG ACC ACC CAC ACC Trp Leu Pro Asn Pro Asn Leu Lys Ala Glu Arg Thr Thr Thr His Thr 530 535 540	1632
CTC TCT CTG CAA GGC CGC AGC GAA AAA GGT ACT TTG GAT GCC AAC CTG Leu Ser Leu Gln Gly Arg Ser Glu Lys Gly Thr Leu Asp Ala Asn Leu 545 550 555 560	1680
TAT CAA AGC AAT TAC CGC AAT TTC CTG TCT GAA GAG CAG AAG CTG ACC Tyr Gln Ser Asn Tyr Arg Asn Phe Leu Ser Glu Glu Gln Lys Leu Thr 565 570 575	1728
ACC AGC GGC GAT GTC AGC TGT ACT CAG ATG AAT TAC TAC TAC GGT ATG Thr Ser Gly Asp Val Ser Cys Thr Gln Met Asn Tyr Tyr Gly Met 580 585 590	1776
TGT AGC AAT CCT TAT TCC GAA AAA CTG GAA TGG CAG ATG CAA AAT ATC Cys Ser Asn Pro Tyr Ser Glu Lys Leu Glu Trp Gln Met Gln Asn Ile 595 600 605	1824
GAC AAG GCC AGA ATC CGC GGT ATC GAG CTG ACG GGC CGT CTG AAT GTG Asp Lys Ala Arg Ile Arg Gly Ile Glu Leu Thr Gly Arg Leu Asn Val 610 615 620	1872
GAC AAA GTA GCG TCT TTT GTT CCT GAG GGC TGG AAA CTG TTC GGC TCG Asp Lys Val Ala Ser Phe Val Pro Glu Gly Trp Lys Leu Phe Gly Ser 625 630 635 640	1920
CTG GGT TAT GCG AAA AGC AAA CTG TCG GGC GAC AAC AGC CTG CTG TCC Leu Gly Tyr Ala Lys Ser Lys Leu Ser Gly Asp Asn Ser Leu Leu Ser 645 650 655	1968
ACC CAG CCG TTG AAA GTG ATT GCC GGT ATC GAC TAT GAA AGT CCG AGC Thr Gln Pro Leu Lys Val Ile Ala Gly Ile Asp Tyr Glu Ser Pro Ser 660 665 670	2016
GAA AAA TGG GGC GTG TTC TCC CGC CTG ACC TAT CTG GGC GCG AAA AAG Glu Lys Trp Gly Val Phe Ser Arg Leu Thr Tyr Leu Gly Ala Lys Lys 675 680 685	2064
GTC AAA GAC GCG CAA TAC ACC GTT TAT GAA AAC AAG GGC TGG GGT ACG Val Lys Asp Ala Gln Tyr Thr Val Tyr Glu Asn Lys Gly Trp Gly Thr 690 695 700	2112
CCT TTG CAG AAA AAG GTA AAA GAT TAC CCG TGG CTG AAC AAG TCG GCT Pro Leu Gln Lys Lys Val Lys Asp Tyr Pro Trp Leu Asn Lys Ser Ala 705 710 715 720	2160
TAT GTG TTC GAT ATG TAC GGC TTC TAC AAA CCG GTG AAA AAC CTG ACT Tyr Val Phe Asp Met Tyr Gly Phe Tyr Lys Pro Val Lys Asn Leu Thr 725 730 735	2208
TTG CGT GCA GGC GTA TAT AAT GTG TTC AAC CGC AAA TAC ACC ACT TGG Leu Arg Ala Gly Val Tyr Asn Val Phe Asn Arg Lys Tyr Thr Thr Trp 740 745 750	2256
GAT TCC CTG CGC GGC CTG TAT AGC TAC AGC ACC ACC AAC TCG GTC GAC Asp Ser Leu Arg Gly Leu Tyr Ser Tyr Ser Thr Thr Asn Ser Val Asp 755 760 765	2304

Figure 8 (cont'd.)

CGC	GAT	GGC	AAA	GGC	TTA	GAC	CGC	TAC	CGC	GCC	CCA	AGC	CGT	AAT	TAC	2352
Arg	Asp	Gly	Lys	Gly	Leu	Asp	Arg	Tyr	Arg	Ala	Pro	Ser	Arg	Asn	Tyr	
770						775					780					
GCC	GTA	TCG	CTG	GAA	TGG	AAG	TTT	TAA								2379
Ala	Val	Ser	Leu	Glu	Trp	Lys	Phe	*								
785						790										

Figure 9

ATG AAA CCA TTA CAC ATG CTT CCT ATT GCC GCG CTG GTC GGC AGT ATT Met Lys Pro Leu His Met Leu Pro Ile Ala Ala Leu Val Gly Ser Ile 1 5 10 15	48
TTC GGC AAT CCG GTC TTG GCA GCG GAT GAA GCT GCA ACC GAA ACC ACA Phe Gly Asn Pro Val Leu Ala Ala Asp Glu Ala Ala Thr Glu Thr Thr 20 25 30	96
CCC GTT AAA GCA GAG ATA AAA GAA GTG CGC GTT AAA GAC CAG CTT AAT Pro Val Lys Ala Glu Ile Lys Glu Val Arg Val Lys Asp Gln Leu Asn 35 40 45	144
GCG CCT GCA ACC GTG GAA CGT GTC AAC CTC GGC CGC ATT CAA CAG GAA Ala Pro Ala Thr Val Glu Arg Val Asn Leu Gly Arg Ile Gln Gln Glu 50 55 60	192
ATG ATA CGC GAC AAC AAA GAC TTG GTG CGT TAC TCC ACC GAC GTC GGC Met Ile Arg Asp Asn Lys Asp Leu Val Arg Tyr Ser Thr Asp Val Gly 65 70 75 80	240
TTG AGC GAT AGC GGC CGC CAT CAA AAA GGC TTT GCT GTG CGC GGC GTG Leu Ser Asp Ser Gly Arg His Gln Lys Gly Phe Ala Val Arg Gly Val 85 90 95	288
GAA GGC AAC CGT GTC GGT GTC AGC ATT GAC GGC GTG AGC CTG CCT GAT Glu Gly Asn Arg Val Gly Val Ser Ile Asp Gly Val Ser Leu Pro Asp 100 105 110	336
TCG GAA GAA AAC TCA CTG TAT GCA CGT TAT GGC AAC TTC AAC AGC TCG Ser Glu Glu Asn Ser Leu Tyr Ala Arg Tyr Gly Asn Phe Asn Ser Ser 115 120 125	384
CGC CTG TCT ATC GAC CCC GAA CTC GTG CGC AAC ATC GAA ATC GCG AAG Arg Leu Ser Ile Asp Pro Glu Leu Val Arg Asn Ile Glu Ile Ala Lys 130 135 140	432
GGC GCT GAC TCT TTC AAT ACC GGT AGC GGC GCA TTG GGT GGC GGC GTG Gly Ala Asp Ser Phe Asn Thr Gly Ser Gly Ala Leu Gly Gly Gly Val 145 150 155 160	480
AAT TAC CAA ACC CTG CAA GGA CAT GAT TTG CTG TTG GAC GAC AGG CAA Asn Tyr Gln Thr Leu Gln Gly His Asp Leu Leu Leu Asp Asp Arg Gln 165 170 175	528
TTC GGC GTG ATG ATG AAA AAC GGT TAC AGC AGC CGC AAC CGC GAA TGG Phe Gly Val Met Met Lys Asn Gly Tyr Ser Ser Arg Asn Arg Glu Trp 180 185 190	576
ACA AAT ACA CTC GGT TTC GGT GTG AGC AAC GAC CGC GTG GAT GCC GCT Thr Asn Thr Leu Gly Phe Gly Val Ser Asn Asp Arg Val Asp Ala Ala 195 200 205	624
TTG CTG TAT TCG CAA CGT CGC GGT CAT GAG ACC GAA AGC GCG GGC GAG Leu Leu Tyr Ser Gln Arg Arg Gly His Glu Thr Glu Ser Ala Gly Glu 210 215 220	672
CGT GGC TAT CCG GTA GAG GGT GCT GGC AGC GGA GCA ATT ATC CGT GGT Arg Gly Tyr Pro Val Glu Gly Ala Gly Ser Gly Ala Ile Ile Arg Gly 225 230 235 240	720
TCG TCA CGC GGT ATC CCT GAT CCG TCC AAA CAC AAA TAC CAC AAC TTC Ser Ser Arg Gly Ile Pro Asp Pro Ser Lys His Lys Tyr His Asn Phe 245 250 255	768

Figure 9 (cont'd.)

TTG GGT AAG ATT GCT TAT CAA ATC AAC GAC AAG CAC CGC ATC GGC CCA Leu Gly Lys Ile Ala Tyr Gln Ile Asn Asp Lys His Arg Ile Gly Pro 260 265 270	816
TCG TTT AAC GGC CAG CAG GGG CAT AAT TAC ACG ATT GAA GAG TCT TAT Ser Phe Asn Gly Gln Gln Gly His Asn Tyr Thr Ile Glu Glu Ser Tyr 275 280 285	864
AAC CTG ACC GCT TCT TCC TGG CGC GAA GCC GAT GAC GTA AAC AGA CGG Asn Leu Thr Ala Ser Ser Trp Arg Glu Ala Asp Asp Val Asn Arg Arg 290 295 300	912
CGC AAT GCC AAC CTC TTT TAC GAA TGG ACG CCT GAT TCA AAT TGG CTG Arg Asn Ala Asn Leu Phe Tyr Glu Trp Thr Pro Asp Ser Asn Trp Leu 305 310 315 320	960
TCG TCT TTG AAG GCG GAC TTC GAT TAT CAG ACA ACC AAA GTG GCG GCG Ser Ser Leu Lys Ala Asp Phe Asp Tyr Gln Thr Thr Lys Val Ala Ala 325 330 335	1008
GTT AAC AAC AAA GGC TCG TTC CCG ACG GAT TAT TCC ACC TGG ACG CGC Val Asn Asn Lys Gly Ser Phe Pro Thr Asp Tyr Ser Thr Trp Thr Arg 340 345 350	1056
AAC TAT AAT CAG AAG GAT TTG GAG AAT ATA TAC AAC CGC AGC ATG GAC Asn Tyr Asn Gln Lys Asp Leu Glu Asn Ile Tyr Asn Arg Ser Met Asp 355 360 365	1104
ACC CGA TTC AAA CGT TTT ACT TTG CGT ATG GAC AGC CAA CCG TTG CAA Thr Arg Phe Lys Arg Phe Thr Leu Arg Met Asp Ser Gln Pro Leu Gln 370 375 380	1152
CTG GGC GGC CAA CAT CGC TTG TCG CTT AAA ACT TTC GCC AGT CGG CGT Leu Gly Gly Gln His Arg Leu Ser Leu Lys Thr Phe Ala Ser Arg Arg 385 390 395 400	1200
GAG TTT GAA AAC TTA AAC CGC GAC GAT TAT TAC TTC AGC GAA AGA GTA Glu Phe Glu Asn Leu Asn Arg Asp Asp Tyr Tyr Phe Ser Glu Arg Val 405 410 415	1248
TCC CGT ACT ACC AGC TCG ATT CAA CAC CCC GTG AAA ACC ACT AAT TAT Ser Arg Thr Thr Ser Ser Ile Gln His Pro Val Lys Thr Thr Asn Tyr 420 425 430	1296
GGT TTC TCA CTG TCT GAT CAA ATC CAA TGG AAC GAC GTG TTC AGC AGC Gly Phe Ser Leu Ser Asp Gln Ile Gln Trp Asn Asp Val Phe Ser Ser 435 440 445	1344
CGT GCA GAT ATC CGT TAC GAT CAT ACC AAA ATG ACG CCT CAG GAA TTG Arg Ala Asp Ile Arg Tyr Asp His Thr Lys Met Thr Pro Gln Glu Leu 450 455 460	1392
AAT GCC GAG TGT CAT GCT TGT GAC AAA ACA CCG CCT GCA GCC AAT ACT Asn Ala Glu Cys His Ala Cys Asp Lys Thr Pro Pro Ala Ala Asn Thr 465 470 475 480	1440
TAT AAA GGC TGG AGC GGA TTT GTC GGT TTG GCG GCG CAA CTG AAT CAG Tyr Lys Gly Trp Ser Gly Phe Val Gly Leu Ala Ala Gln Leu Asn Gln 485 490 495	1488
GCT TGG CAT GTC GGT TAC GAC ATT ACT TCC GGC TAC CGT GTC CCC AAT Ala Trp His Val Gly Tyr Asp Ile Thr Ser Gly Tyr Arg Val Pro Asn 500 505 510	1536

Figure 9 (cont'd.)

GCG TCC GAA GTG TAT TTC ACT TAC AAC CAC GGT TCG GGT AAT TGG CTG Ala Ser Glu Val Tyr Phe Thr Tyr Asn His Gly Ser Gly Asn Trp Leu 515 520 525	1584
CCC AAT CCC AAC CTG AAA GCC GAG CGC AGC ACC ACC CAC ACC CTG TCT Pro Asn Pro Asn Leu Lys Ala Glu Arg Ser Thr Thr His Thr Leu Ser 530 535 540	1632
CTG CAA GGC CGC AGC GAA AAA GGT ACT TTG GAT GCC AAC CTG TAT CAA Leu Gln Gly Arg Ser Glu Lys Gly Thr Leu Asp Ala Asn Leu Tyr Gln 545 550 555 560	1680
AAC AAT TAC CGC AAC TTC TTG TCT GAA GAG CAG AAG CTG ACC ACC AGC Asn Asn Tyr Arg Asn Phe Leu Ser Glu Glu Gln Lys Leu Thr Thr Ser 565 570 575	1728
GGC GAT GTC GGC TGT ACT CAG ATG AAT TAC TAC TAC GGT ATG TGT AGC Gly Asp Val Gly Cys Thr Gln Met Asn Tyr Tyr Tyr Gly Met Cys Ser 580 585 590	1776
AAT CCT TAT TCC GAA AAA CCG GAA TGG CAG ATG CAA AAT ATC GAT AAG Asn Pro Tyr Ser Glu Lys Pro Glu Trp Gln Met Gln Asn Ile Asp Lys 595 600 605	1824
GCC CGA ATC CGT GGT CTT GAG CTG ACA GGC CGT CTG AAT GTG ACA AAA Ala Arg Ile Arg Gly Leu Glu Leu Thr Gly Arg Leu Asn Val Thr Lys 610 615 620	1872
GTA GCG TCT TTT GTT CCT GAG GGC TGG AAA TTG TTC GGC TCG CTG GGT Val Ala Ser Phe Val Pro Glu Gly Trp Lys Leu Phe Gly Ser Leu Gly 625 630 635 640	1920
TAT GCG AAA AGC AAA CTG TCG GGC GAC AAC AGC CTG CTG TCC ACA CAG Tyr Ala Lys Ser Lys Leu Ser Gly Asp Asn Ser Leu Leu Ser Thr Gln 645 650 655	1968
CCG CCG AAA GTG ATT GCC GGT GTC GAC TAC GAA AGC CCG AGC GAA AAA Pro Pro Lys Val Ile Ala Gly Val Asp Tyr Glu Ser Pro Ser Glu Lys 660 665 670	2016
TGG GGT GTG TTC TCC CGC CTG ACT TAT CTG GGT GCG AAA AAG GCC AAA Trp Gly Val Phe Ser Arg Leu Thr Tyr Leu Gly Ala Lys Lys Ala Lys 675 680 685	2064
GAC GCG CAA TAC ACC GTT TAT GAA AAC AAG GGC CGG GGT ACG CCT TTG Asp Ala Gln Tyr Thr Val Tyr Glu Asn Lys Gly Arg Gly Thr Pro Leu 690 695 700	2112
CAG AAA AAG GTA AAA GAT TAC CCG TGG CTG AAC AAG TCG GCT TAT GTG Gln Lys Lys Val Lys Asp Tyr Pro Trp Leu Asn Lys Ser Ala Tyr Val 705 710 715 720	2160
TTT GAT ATG TAC GGC TTC TAC AAA CTG GCT AAA AAC CTG ACT TTG CGT Phe Asp Met Tyr Gly Phe Tyr Lys Leu Ala Lys Asn Leu Thr Leu Arg 725 730 735	2208
GCA GGC GTA TAT AAT GTG TTC AAC CGC AAA TAC ACC ACT TGG GAT TCC Ala Gly Val Tyr Asn Val Phe Asn Arg Lys Tyr Thr Thr Trp Asp Ser 740 745 750	2256
CTG CGC GGT TTG TAT AGC TAC AGC ACC ACC AAC GCG GTC GAC CGA GAT Leu Arg Gly Leu Tyr Ser Tyr Ser Thr Thr Asn Ala Val Asp Arg Asp 755 760 765	2304



Figure 9 (cont'd.)

GGC	AAA	GGC	TTA	GAC	CGC	TAC	CGC	GCC	TCA	GGC	CGT	AAT	TAC	GCC	GTA	2352
Gly	Lys	Gly	Leu	Asp	Arg	Tyr	Arg	Ala	Ser	Gly	Arg	Asn	Tyr	Ala	Val	
	770					775					780					
TCG	CTG	GAT	TGG	AAG	TTT	TGA	ATTCC									2378
Ser	Leu	Asp	Trp	Lys	Phe	*										
785					790											

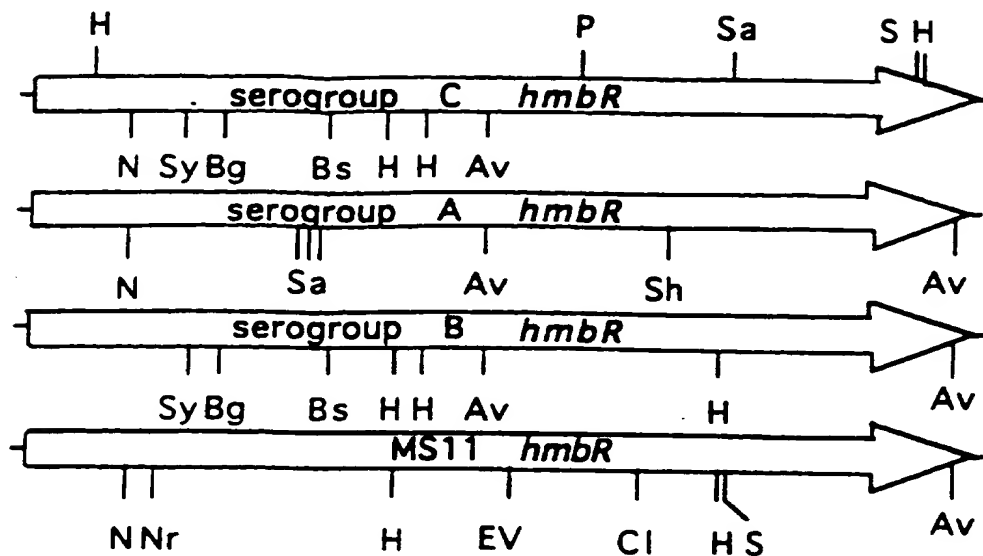


Fig. 10

## FIGURE 11

HMBRA	MKPLQMLPIAALVGSIFGNPVLAADAAATETTPVKAIEIKAVRVKGQRNAP	50
HMBRB	MKPLQMPPIAALLGSIFGNPVFAADAAATETTPVKAIEIKAVRVKGQRNAP	50
HMBRC	MKPLQMLPIAALVGSIFGNPVFAADAAATETTPVKAIEIKAVRVKGQRNAP	50
HMBRMS11	MKPLHMLPIAALVGSIFGNPVLAADAAATETTPVKAIEIKAVRVKQDLNAP *****	50
HMBRA	AAVERVNLNRIKQEMIRDNDLVRYSTDVGLSDSGRHQKGFVRGVEGDR	100
HMBRB	AAVERVNLNRIKQEMIRDNDLVRYSTDVGLSDSRHQKGFVRGVEGDR	100
HMBRC	AAVERVNLNRIKQEMIRDNDLVRYSTDVGLSDSGRHQKGFVRGVEGDR	100
HMBRMS11	ATVERVNLGRIQQEMIRDNDLVRYSTDVGLSDSGRHQKGFVRGVEGDR *****	100
HMBRA	VGVSIDGVNLPDSEENSLYARYGNFNSSRLSIDPELVNIEIVKGADSFN	150
HMBRB	VGVSIDGVNLPDSEENSLYARYGNFNSSRLSIDPELVNIDIVKGADSFN	150
HMBRC	VGVSIDGVNLPDSEENSLYARYGNFNSSRLSIDPELVNIDIVKGADSFN	150
HMBRMS11	VGVSIDGVSLPDSEENSLYARYGNFNSSRLSIDPELVNIEIAKGADSFN *****	150
HMBRA	TGSGALGGGVNYQTLOGRDLLLDDRQFGVMMKNGYSTRNREWTNTLGFGV	200
HMBRB	TGSGALGGGVNYQTLOGRDLLLPERQFGVMMKNGYSTRNREWTNTLGFGV	200
HMBRC	TGSGALGGGVNYQTLOGRDLLLPERQFGVMMKNGYSTRNREWTNTLGFGV	200
HMBRMS11	TGSGALGGGVNYQTLOGHDLDDLDDRQFGVMMKNGYSTRNREWTNTLGFGV *****	200
HMBRA	SNDRVDAALLYSQRRGHETESAGNRGYPVEGAGKETNIRGSARGIPDPSK	250
HMBRB	SNDRVDAALLYSQRRGHETESAGKRGYPVEGAGSGANIRGSARGIPDPSQ	250
HMBRC	SNDRVDAALLYSQRRGHETESAGKRGYPVEGAGSGANIRGSARGIPDPSQ	250
HMBRMS11	SNDRVDAALLYSQRRGHETESAGERGYVEGAGSGAIRGSSRGIPDPSK *****	250
HMBRA	HKYHNFGLKIAIYQINDNHRIGASLNGQQGHNYTVEESYNLTASSWREADD	300
HMBRB	HKYHSFLGKIAIYQINDNHRIGASLNGQQGHNYTVEESYNLLASYWREADD	300
HMBRC	HKYHSFLGKIAIYQINDNHRIGASLNGQQGHNYTVEESYNLLASYWREADD	300
HMBRMS11	HKYHNFGLKIAIYQINDKHRIGPSFNGQQGHNYTVEESYNLTASSWREADD *****	300
HMBRA	VNRRRNANLFYEWMPDSNWLSSLKADFDYQKTKVAAIN-KGSPPT-NYTT	348
HMBRB	VNRRRNTNLFYEWTPESDRLSMVKADVDYQKTKVSAVNYKGSFPT-NYTT	349
HMBRC	VNRRRNTNLFYEWTPESDRLSMVKADVDYQKTKVSAVNYKGSFPEDSST	350
HMBRMS11	VNRRRNANLFYEWTPDSNWLSSLKADFDYQKTKVAAVNNKGSFPTD-YST *****	349
HMBRA	WETEHKKEVGEIYNRSMDFRFRITLRLDSHPLQLGGGRHRLSFKTFAS	398
HMBRB	WETEHKKEVGEIYNRSMDFRFRITLRLDSHPLQLGGGRHRLSFKTFAG	399
HMBRC	LTRNYNQDLDEIYNRSMDFRFRITLRLDSHPLQLGGGRHRLSFKTFAS	400
HMBRMS11	WTRNYNQDLNENIYNRSMDFRFRITLRLDSHPLQLGG-QHRLSLKTFAS *****	398
HMBRA	RRDFENLNRDDYYFSGRVVRTTSSIQHPVKTTNYGFSLSQIQWNDVFSS	448
HMBRB	QRDFENLNRDDYYFSGRVVRTTNSIQHPVKTTNYGFSLSQIQWNDVFSS	449
HMBRC	RRDFENLNRDDYYFSGRVVRTTSSIQHPVKTTNYGFSLSQIQWNDVFSS	450
HMBRMS11	RREFENLNRDDYYFSERSVRTTSSIQHPVKTTNYGFSLSQIQWNDVFSS *****	448
HMBRA	RAGIRYDHTKMTPOELNAECHACDKTPPAANTYKGWSGFVGLAAQLNQAW	498
HMBRB	RAGIRYDHTKMTPOELNADCHACDKTPPAANTYKGWSGFVGLAAQLSQTW	499
HMBRC	RAGIRYDHTKMTPOELNAECHACDKTPPAANTYKGWSGFVGLAAQLNQAW	500

HMBRMS11	RAIIRYDHTYKMTPOELNAECHACDKTPPAANTYKGWSGFVGLAAQLNQAW *****	498
HMBRA	RVGYDITSGYRVPNASEVYFTYNHGSGNWL PNP NLKAERSTTHTLSLQGR	548
HMBRB	RLGYDVTSGFRVPNASEVYFTYNHGSGTWKPNP NLKAERSTTHTLSLQGR	549
HMBRC	RVGYDITSGYRVPNASEVYFTYNHGSGNWL PNP NLKAERTTHTLSLQGR	550
HMBRMS11	HVGYDITSGYRVPNASEVYFTYNHGSGNWL PNP NLKAERSTTHTLSLQGR *****	548
HMBRA	SEKGM LDANLYQSNYRNFLSEEQKLTTSGTPGCTEENAYYSICSDDPYKEK	598
HMBRB	GDKGTLDANLYQSNYRNFLSEEQNLTVSGTPGCTEEDAYYYRCSDPYKEK	599
HMBRC	SEKGTLDANLYQSNYRNFLSEEQKLTTSGDV SCTQMNYYYGMC SNPYSEK	600
HMBRMS11	SEKGTLDANLYQSNYRNFLSEEQKLTTSGDV GCTQMNYYYGMC SNPYSEK *****	598
HMBRA	LDWQMKNI DKARIRGIELTGRLNVDKVASFVPEGWKLF GSLGYAKSKLSG	648
HMBRB	LDWQMKNI DKARIRGIELTGRLNVDKVASFVPEGWKLF GSLGYAKSKLSG	649
HMBRC	LEWQMNI DKARIRGIELTGRLNVDKVASFVPEGWKLF GSLGYAKSKLSG	650
HMBRMS11	PEWQMNI DKARIRGLELTGRLNVT KVASFVPEGWKLF GSLGYAKSKLSG *****	648
HMBRA	DNSLLSTQPLKVIAGIDYESPSEKGVFSRLTYLGAKKVKDAQYTVYENK	698
HMBRB	DNSLLSTQPLKVIAGIDYESPSEKGVFSRLTYLGAKKVKDAQYTVYENK	699
HMBRC	DNSLLSTQPLKVIAGIDYESPSEKGVFSRLTYLGAKKVKDAQYTVYENK	700
HMBRMS11	DNSLLSTQPPKVIAGVDYESPSEKGVFSRLTYLGAKKAKDAQYTVYENK *****	698
HMBRA	GWGTPLQKKVKDYPWL NKSAYVFD MYGFYKPVKNLT L RAGVYNLFNRKYT	748
HMBRB	GWGTPLQKKVKDYPWL NKSAYVFD MYGFYKPAKNLT L RAGVYNLFNRKYT	749
HMBRC	GWGTPLQKKVKDYPWL NKSAYVFD MYGFYKPVKNLT L RAGVYNVFNRYT	750
HMBRMS11	GRGTPLQKKVKDYPWL NKSAYVFD MYGFYKLAKNLT L RAGVYNVFNRYT *****	748
HMBRA	TWDSL RGLYSYSTTNAVD RDGKGLDRYRAPGRNYAVSLEWKF	790
HMBRB	TWDSL RGLYSYSTTNAVD RDGKGLDRYRAPGRNYAVSLEWKF	791
HMBRC	TWDSL RGLYSYSTTNSVDRD GKGLDRYRAPSRNYAVSLEWKF	792
HMBRMS11	TWDSL RGLYSYSTTNAVD RDGKGLDRYRASGRNYAVSLDWKF *****	790

Identity : 671 ( 84.7%)

Similarity: 92 ( 11.6%)

FIGURE 11 (cont'd.)